SUBSTRATES, STRUCTURES, AND FUNCTIONS OF THE CHLOROPLAST CLP PROTEASE SYSTEM IN *ARABIDOPSIS THALIANA*

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Jui-Yun Liao

May 2019

SUBSTRATES, STRUCTURES, AND FUNCTIONS OF THE CHLOROPLAST CLP PROTEASE SYSTEM IN *ARABIDOPSIS THALIANA*

Jui-Yun Liao, Ph. D.
Cornell University 2019

The caseinolytic proteolytic machinery (CLP) is an essential and abundant protease of the chloroplast protease network. It is composed of multiple components (a proteolytic core CLPP/R/T, chaperones CLPC1/C2/D, and adaptors CLPS1/F). Mostly based on functional and structural information from bacterial Clp systems, it is postulated that these chloroplast CLP chaperones are aided by the CLP adaptors to select and deliver substrates to the proteolytic chamber (protease core) for degradation. The chloroplast CLPPRT proteolytic core is different and far more complex than the bacterial or mitochondria Clp core. The chloroplast CLP core is a hetero-oligomeric tetradecamer that is associated with additional accessory proteins unique to higher plants. Furthermore, the chloroplast CLPP and CLPR subunits have C-terminal extensions with unknown functions. It is unclear why chloroplast CLP core shows such high complexity and how these different CLP subunits contribute to the proteolytic system. Finally, relatively few chloroplast CLP substrates have been identified.

To better understand the chloroplast CLP protease system, I applied an *in vivo* trapping approach for substrate identification and crosslinking (XL) mass

spectrometry (MS) for investigation of the proximity and possible protein-protein interactions between these CLP components. Functional complementation showed that CLPP5 is crucial for CLP catalysis, whereas CLPP3 plays an essential role in CLP structure but its catalytic activity is dispensable. However, *in vivo* trapping using CLPPRT complexes with a reduced number of catalytic triads through the presence of one or more catalytically inactivated CLPP3/5 subunits did not identify proteins trapped in these CLPPRT complexes. This suggests that reduced proteolytic capacity within CLP cores does not result in a bottleneck for protein degradation *in vivo*. XL-MS of affinity-purified CLP core complexes or affinity purified CLPC-TRAP complexes identified several putative domains and motifs involved in the CLP protein-protein interactions. The newly established workflow of *in vitro* DSSO crosslinking using plant proteins paves the way for a more detailed exploration of the 3D structure and possible regulation of the chloroplast CLP machinery.

BIOGRAPHICAL SKETCH

Born and raised in Taipei, Taiwan, Jui-Yun Rei Liao graduated from National Chung Hsing University with a Bachelor of Science degree in Plant Pathology in 2007. She completed her Master of Science degree in Plant Biology at National Taiwan University in 2010, studying gene regulation of a protease inhibitor in sweet potatoes. In 2013, she entered the graduate program in Plant Biology at Cornell University, where she has worked in the laboratory of Dr. Klaas J. van Wijk. Outside of the laboratory, Jui-Yun enjoys food, novels, movies, jazz, and jogging.

ACKNOWLEDGMENTS

I thank Dr. Klaas J. van Wijk for providing me a wonderful opportunity to conduct my PhD studies in his lab. I appreciate his patience and advice during my graduate studies as well as his encouraging and rigorous approach toward scientific research. This really makes me enjoy working on challenging projects and strongly appreciate the privilege of conducting research.

Second, I thank all van Wijk lab members for their generous support and help during my PhD journey, especially Dr. Giulia Friso for her caring disposition and help in MS/MS analysis as well as Dr. Jitae Kim for his guidance in molecular biology and proteome techniques.

Third, I thank my committee members Dr. Shu-Bing Qian and Dr. Maureen Hanson for their time and constructive feedback on the dissertation projects. I also thank Dr. Lan Huang and Dr. Clinton Yu at UC Irvine for their help analyzing the XL-MS samples.

Last, I thank my family and friends, especially Yu-Husan Lin, Yu-Ling Tai, and Xu Lan for their constant support and companionship during my graduate studies.

Financial support was provided by NSF (MCB#1614629 to K.J.v.W.), Cornell TAship, and study abroad fellowship from Ministry of Education in Taiwan.

TABLE OF CONTENTS

Page #

Biographical Sketch.....iii

Acknowledgements……iv

Table of Contents……v

List of Figures……vi

List of Tables……...viii

Chapter One Introduction…….....1

Chapter Two Discovery of AAA+ protease substrates through trapping approaches…….39

Chapter Three Consequences of the loss of catalytic triads in chloroplast CLPPR protease core complexes *in vivo*……75

Chapter Four Investigation of the spatial proximity of the CLP subunits within the chloroplast CLP machinery……124

Chapter Five Conclusions and future perspective……149

LIST OF FIGURES

Page #

CHAPTER ONE

Figure 1. Chloroplast biogenesis and plastids……5

Figure 2. Proteases and plastids……...14

Figure 3. The architecture, components, and current model of CLP proteases……...20

CHAPTER TWO

Figure 1. Schematic outline of *in vivo* substrate traps……..50

Figure 2. The workflow of *in vivo* substrate trapping…….....51

CHAPTER THREE

Figure 1. Characterization of *clpp3-1* complemented with *CLPP3-STREPll* and *CLPP3S164A-STREPll*…….82

Figure 2. Characterization of wt and *clpp5-1* lines expressing *CLPP5S193A-STREPII* or *CLPP5-STREPll* transgenes……...92

Figure 3. Affinity purification of STREPII-tagged CLP core complexes and MS/MS analysis……98

Supplemental Figure 1. Sequence alignment of CLPP3 and CLPP5 with CLPP homologs…….....84

Supplemental Figure 2. Genotyping and RT-PCR analysis of transgenic lines…….85

Supplemental Figure 3. MS/MS-based verification of the S164A mutation in CLPP3S164A-STREPII and comparison to CLPP3-STREPII…….88

Supplemental Figure 4. MS/MS-based verification of the S193A mutation in CLPP5S193A-STREPII and comparison to CLPP5-WT-STREPII…….....94

Supplemental Figure 5. Effect of preincubation of soluble leaf proteomes with avidin to reduce the binding and enrichment on streptactin columns for CLPP3(S164A)-STREPII and CLPP5(S193A)-STREPII lines……..100

CHAPTER FOUR

Figure 1. The workflow of XL-MS……128

Figure 2. Western blot of the affinity eluates before and after DSSO crosslinking.....132

Figure 3. Proximity matrix of the CLPPRT subunits according to the intersubunit links…….....135

Figure 4. The crystal structure of the protomer of *E. coli* CLPP 1YG6…….137

LIST OF TABLES

Page #

CHAPTER TWO

Table 1. Overview and summary of highlights and key information of published substrate trapping studies for the AAA+ proteolytic systems of Clp, FtsH and Lon families…….....43

CHAPTER THREE

Supplemental Table 1. Primer sets used in P3-P5 study……....87

Supplemental Table 2. Proteins identified across the 16 affinity experiments with CLPP3-STREPII, CLPP3S164A-STREPII, CLPP5-STREPII or CLPP5S193A-STREPll.…….105

CHAPTER FOUR

Table 1. Summary of the interlinks among CLP subunits identified through XL-MS of the eluates from strep-tagged CLPP5/R4 lines…….....135

Table 2. Projection of the lysine residue on high resolution crystal structure EcCLPP (1yg6)……...137

CHAPTER ONE

INTRODUCTION

1.1 OVERVIEW OF PLASTID BIOGENESIS

1.1.1 Plastid types and functions in higher plants

Plastids are a family of cellular organelles that originate from a photosynthetic bacterium through endosymbiosis (*1, 2*). Plastids contain a relatively small genome encoding for ribosomal RNAs, various proteins of the plastid gene expression machinery, proteins involved in photosynthesis, and small number of proteins involved with metabolism (*e.g.* fatty acids) and biogenesis (*e.g.* Clp protease).

Higher plants contain several types of several plastid types, all originating from proplastids. Plastid types include pro-plastids, etioplasts, amyloplasts, chromoplasts, chloroplasts, leucoplasts, and gerontoplasts (Figure 1A). These various types of plastids are characterized according to their unique pigments or storage molecules within them. The pro-plastid is the progenitor of the plastid family. It is present in the meristem and in undifferentiated, non-photosynthetic dividing cells such as developing seeds. Proplastids lack the pigments chlorophyll and carotenoids, and have few poorly organized thylakoid membranes and possible starch grains. The leucoplast, which has no chlorophyll but can have some carotenoids, is found in non-photosynthetic tissues such as roots or white petals. It stores organic compounds such as proteins, lipids, and starch. The amyloplast is characterized by its storage starch grains and is found in storage tissues such as seed endosperm (in particular in grasses),

and tuberous roots (*e.g.* potato, cassava). In root tips, the sedimentation of the amyloplasts in columella cells triggers signaling and regulates gravitropism. The chromoplast is a pigmented organelle highly enriched with carotenoids stored in plastoglobules (globular or fibrillous) which give fruits and flowers color and flavor. The chloroplast is present in all photosynthetic tissues and is the most studied plastid type due to its role in photosynthesis. Illumination triggers differentiation of proplastids and etioplasts into chloroplasts. The proplastid will differentiate into pre-granal plastid and then differentiate into the mature chloroplast. During this differentiation pathway, thylakoid membranes are made and organized in the differentiating plastid. Without light, plastids in leaves develop into etioplasts, but chloroplasts or pre-granal plastids can also de-differentiate into etioplasts if moved into darkness for an extended period of time. The etioplast contains a crystalline prolamellae body that includes protochlorophyllide reductase as well as proto-chlorophyllide. Upon exposure to light, etioplasts rapidly (within 24 hours) differentiate into the chloroplasts. During senescence, chloroplasts differentiate into gerontoplasts, which have strongly reduced thylakoid and stromal protein content. This chloroplast-to-gerontoplast transition recycles the nutrients in chloroplasts, and it is part of the cell senescence program (*3-8*). Most plastid types can differentiate or de-differentiate to other plastid types during cell growth and development, and in response to environmental conditions.

1.1.2 Chloroplast biogenesis and regulation

The chloroplast is specialized in photosynthesis, as well as a number of important metabolic functions and has an extensive thylakoid membrane system. The

outer boundary of the chloroplast, and all other plastid types, is a double envelope membrane system. The outer envelope membrane, which likely originated from the engulfed prokaryotes (*9*), separates the chloroplast from the cytosol. The inner membrane, which evolved from the membrane of the engulfed cyanobacteria, separates the inter-envelope space and the stroma. Both outer and inner envelope membranes are lipid bilayers and contain transporters and protein complexes for the exchange of ions and molecules between chloroplasts and cytosol. The thylakoid membrane is likely generated (at least in part) from the chloroplast inner envelope membrane, suggested by the observation of small vesicles budding from the inner envelope, and the fact that thylakoid membrane has a similar lipid composition as the chloroplast inner membrane (*10*). The thylakoid membrane is the site for the light-dependent reactions of photosynthesis and is organized in grana stacks and stroma lamella. The outer leaflet (facing the stroma) of the thylakoid membrane can form lipoprotein particles, named plastoglobules; these are particularly important during abiotic stress, and during developmental transitions such as senescence (*11*). Enclosed by the thylakoid membranes is the thylakoid lumen with a low pH (3.5-7), driven by proton translocation across the thylakoid during light-driven photosynthetic electron transport. Between the chloroplast envelope and the thylakoid membranes is the stromal region containing the majority of (soluble) proteins including the Calvin-Benson cycle and many metabolic pathways, as well as the 70S ribosomes and most of the gene expression machinery. The chloroplast (and other plastid types) chromosome and its associated proteins (named the nucleoid) is found in the stroma, as well as associated with the inner envelope and thylakoid membrane (Figure 1A-B). Overall,

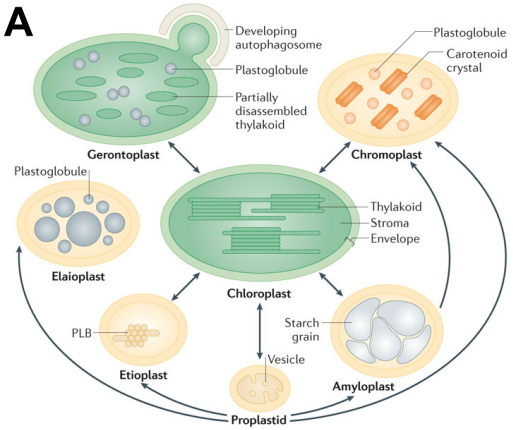
the inner architecture divides the chloroplast into several compartments that allow various biochemical/metabolic reactions occur within the organelle (*3, 8*).

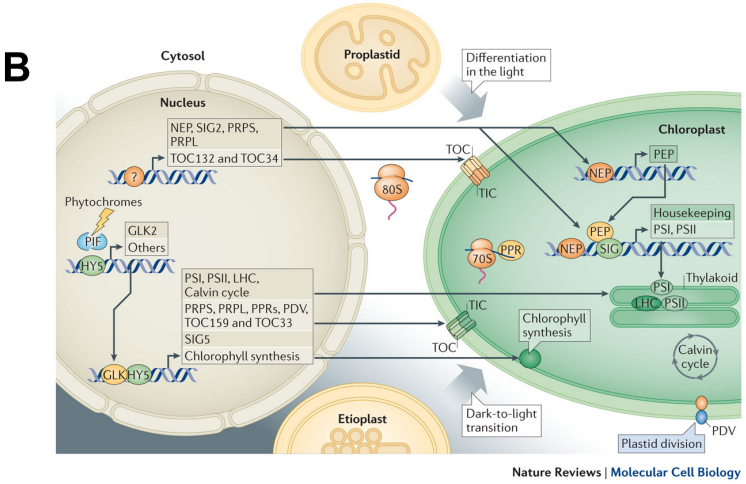
The regulation of chloroplast biogenesis in plants is controlled at both the cellular and genetic level in response to developmental and environmental cues. Light is the key factor to initiate proplastids entering the chloroplast differentiation pathway in dividing cells (meristem or embryo). Upon illumination, the light-activated genes in both the plastid and nucleus genomes are expressed and coordinate to generate thylakoid membranes and photosynthesis machineries. Concurrently, plastids in the dividing cells also propagate and divide during cell division to ensure every daughter cell contains plastids in that cell lineage. This plastid differentiation is cell type-specific. For example, mesophyll cells derived from the L2/L3 layer of meristem have abundant chloroplasts, while epidermal cells from the L1 layer of meristem have only a few chloroplasts. In addition, chloroplast number per cell is regulated by growth and developmental stage. Mature leaf cells often have more chloroplasts than younger leaf cells. During leaf senescence, chloroplast number per cell decreases and chloroplasts differentiate to gerontoplasts by programmed degradation, that includes both extra-plastidic and intra-plastid pathways. The extra-plastidic pathway (both dependent and independent of the autophagy machinery) removes bulk of contents from the plastid into vesicles, while intra-plastid pathways (*e.g.* proteolysis by intra-chloroplast proteases) break down the unwanted molecules within the plastid (*3, 8*) (Figure 1B).

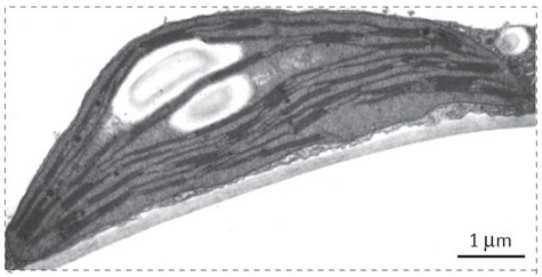
1.1.2.1 Nucleus-encoded chloroplast proteins (focus on import and processing)

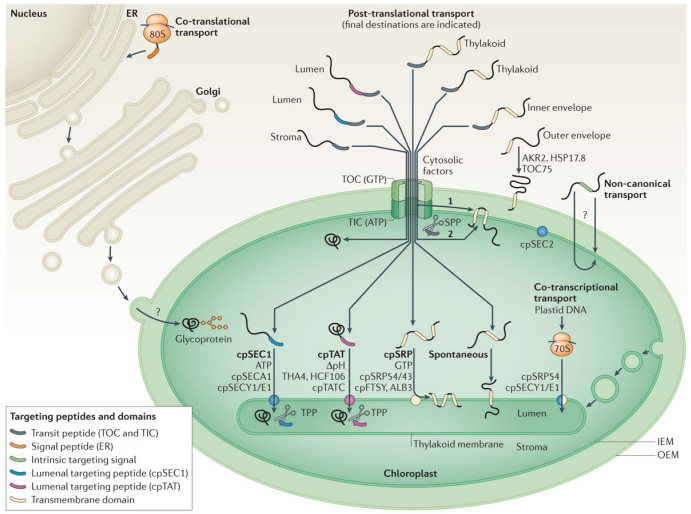
During evolution, most chloroplast genes transferred and integrated into the plant nuclear genome. These transferred genes, adapted to the eukaryotic expression

Figure 1. Plastids and chloroplast biogenesis (*3*). A, Chloroplast Biogenesis and different plastid types. B, Coordination of plastid and nuclear genome during chloroplast biogenesis. C, Chloroplast protein import pathways. Reprinted by permission from Springer Nature Customer Service Centre GmbH: Springer Nature, Nature Review Molecular Cell Biology, Biogenesis and homeostasis of chloroplasts and other plastids. Paul Jarvis and Enrique Lopez-Juez. Copyright © 2014.


A
Developing autophagosome
Plastoglobule
Partially disassembled thylakoid
Plastoglobule
Carotenoid crystal
Gerontoplast
Chromoplast
Plastoglobule
Thylakoid
Stroma
Envelope
Elaioplast
Chloroplast
PLB
Starch grain
Vesicle
Etioplast
Amyloplast
Proplastid


B
Cytosol
Proplastid
Differentiation in the light
Nucleus
NEP, SIG2, PRPS, PRPL
TOC132 and TOC34
TOC
TIC
80S
Chloroplast
PEP
NEP
Phytochromes
PIF
GLK2
Others
HY5
PEP
NEP
SIG
Housekeeping PSI, PSII
70S
PPR
Thylakoid
PSI
LHC
PSII
PSI, PSII, LHC, Calvin cycle
PRPS, PRPL, PPRs, PDV, TOC159 and TOC33
SIG5
Chlorophyll synthesis
TIC
TOC
Chlorophyll synthesis
Calvin cycle
GLK
HY5
Etioplast
Dark-to-light transition
Plastid division
PDV
Nature Reviews | Molecular Cell Biology


1 µm


C
Nucleus
ER
80S
Co-translational transport
Post-translational transport
(final destinations are indicated)
Golgi
Thylakoid
Thylakoid
Lumen
Lumen
Inner envelope
Stroma
Outer envelope
Cytosolic factors
AKR2, HSP17.8
TOC75
TOC (GTP)
Non-canonical transport
1
TIC (ATP)
SPP
2
cpSEC2
?
Co-transcriptional transport
Plastid DNA
?
Glycoprotein
cpSEC1
ATP
cpSECA1
cpSECY1/E1
cpTAT
ΔpH
THA4, HCF106
cpTATC
cpSRP
GTP
cpSRP54/43
cpFTSY, ALB3
Spontaneous
70S
cpSRP54
cpSECY1/E1
TPP
TPP
Lumen
Thylakoid membrane
Stroma
IEM
OEM
Chloroplast
Targeting peptides and domains
Transit peptide (TOC and TIC)
Signal peptide (ER)
Intrinsic targeting signal
Lumenal targeting peptide (cpSEC1)
Lumenal targeting peptide (cpTAT)
Transmembrane domain

system, nuclear genome. These transferred genes, adapted to the eukaryotic expression system, replicate and transcribe in the nucleus and translate in the cytosol, followed by post-translational targeting and import into the chloroplast. Around 3000 chloroplast proteins are nucleus-encoded and targeted to the chloroplast (*3*).

Proteins are targeted to different compartments within the chloroplast through parallel pathways (Figure 1C). Chloroplast outer membrane proteins directly insert into the outer membrane from the cytosol. Proteins targeted to the other sub-compartments within the chloroplast need to cross both envelope membranes and require a signal peptide to guide them to their final destination. The majority of these chloroplast targeted proteins are directed by the N-terminal chloroplast transit peptide (cTP). cTPs are variable in length (20-120 aa) and have very little secondary structure (they are mostly random coils) and lack a consensus sequence. However, cTPs generally do have several features that allow them to be recognized by specifically trained predictors such as TargetP. Some cTPs can be phosphorylated by serine/threonine kinase(s) and recognized by a 'guidance complex' (including at least one Hsp70 and one 14-3-3 protein) in the cytosol. This recognition enhances the transfer of chloroplast precursor proteins to the translocons at the chloroplast outer envelope. However, these precursors have to be dephosphorylated before import (*12, 13*).

Protein translocation across the chloroplast envelope is selective and requires energy in the form of ATP and GTP. The general import pathway contains translocons at the outer or inner chloroplast envelope membranes (TOC or TIC complexes). These translocons contain multiple components and function as a precursor protein receptor

and translocase during protein import. Chaperones in the cytosol (Hsp70) or stroma (Hsp60/70/90/100) are often recruited to assist protein import by preventing the immature folding of precursor proteins after synthesis and hydrolyzing ATP/GTP to provide energy for protein translocation (*3*).

Once precursor proteins enter the stroma, their cTPs are cleaved by the essential stromal processing peptidase (SPP). The cleaved precursor proteins may undergo additional trimming by other peptidases or be re-directed to their final destination based on the emerging signal peptide after SPP cleavage. In particular, lumenal targeting peptides guide proteins to the lumen through the cpSEC1 (chloroplast SEC1) or cpTAT (chloroplast twin arginine translocase) pathway. Similar to cTPs cleaved by SPP, the lumenal targeting peptides are cleaved by thylakoid processing peptidase (TPP or Plsp) and may undergo additional processing by other proteases for full maturation (*3, 14*).

1.1.2.2 Plastid-encoded proteins

The chloroplast genome contains around 100 genes and they can be divided into three groups: i) genes encoding for the gene expression machinery (RNA polymerase, ribosome, rRNA and some tRNA), ii) genes encoding components of photosynthetic machineries (PSI, PSII, cytochrome b6f complex, ATP synthase, and the large subunit of Rubisco), iii) genes encoding for proteins involved in metabolism or biogenesis (CLP, AccD, Ycf1/2).

The plastid transcriptional and translational apparatus shares many similar features with prokaryotes, including prokaryote-like promoters, stable uncapped non-polyA mRNAs, and bacteria-like ribosomes (70S). However, plastid gene expression

needs basal amount of nucleus-encoded components (see 1.1.2.3) and also requires extensive post-transcriptional modifications such as RNA editing (*e.g.* A<u>C</u>U to A<u>U</u>G) and processing of the polycistronic primary transcripts (*3, 15*).

Plastid-encoded proteins tend to be further processed by methionine deformylase (PDF) followed by methionine aminopeptidases (MAP1, MAP2) (*14*). Several plastid-encoded thylakoid membrane proteins co-translationally insert to thylakoid membranes aided by cpSRP (chloroplast signal recognition particle) (*16*). Similar to nucleus-encoded proteins, some plastid-encoded proteins under additional processing. An example is the D1 protein of PSII which only becomes fully functional after its C-terminal processing by the C-terminal processing protease (CtpA) in the lumen (*14*).

1.1.2.3 Coordination of plastid and nuclear-encoded protein biogenesis

Chloroplast protein biogenesis is mediated by light and requires the coordinated expression of plastid- and nucleus-encoded proteins. During the transition of proplastid-to-chloroplast, the nuclear genome expresses a basal amount of RNA polymerase (NEP) and translocons for chloroplast protein import. NEP activates the expression of the plastid-encoded RNA polymerase (PEP). PEP and NEP further work together with PPR (pentatricopeptide repeat) proteins and sigma factor and trigger plastid gene expression. To import a large amount of the proteins (such as the components of LHC, PSI, PSII, and the Calvin cycle), unique photosynthetic translocons encoded by the nuclear genome are formed. In addition, proteins involved in chloroplast protein homeostasis and metabolism are also imported. This type of the coordination where plastid gene expression is under control of the nuclear genome is

known as anterograde signaling. In contrast, plastids also regulate nuclear gene expression through retrograde signaling (*3*) (Figure 1B).

Retrograde signaling was first noted in barley mutants with undifferentiated plastids that had a lower level of proteins encoded by the photosynthesis-associated nuclear genes (*17*). Retrograde signaling can be triggered by genetic defects in the plastid genome, chemicals or inhibitors blocking protein synthesis or metabolic pathways in plastids. Various types of nuclear genes (from photosynthesis-associated to stress-induced genes) can be regulated through retrograde signaling. So far several retrograde pathways have been identified, including the tetrapyrrole intermediate signaling pathway, plastid gene expression pathway, plastid redox state and the reactive oxygen species pathway. These diverse pathways seem to be convergent at some level. GUN1 (genome uncoupling 1) has been proposed as a "master switch" that integrates several retrograde signals within plastids. One model for plastid-nucleus communication is the GUN1-PTM-ABI4 pathway. GUN1 triggers PTM (a plastid envelope-bound protein homeodomain transcription factor) to be cleaved by an unknown protease, resulting in the release of an N-terminal portion of PTM (N-PTM) from the chloroplast envelope to the nucleus. N-PTM interacts with transcriptional factor ABI4, which regulates nuclear gene expression. Whether other routes and mechanisms are involved in plastid-to-nucleus communication is not known (*17-20*). Overall, nuclear and plastid gene expression are closely coordinated for plastid biogenesis and optimization based on development stage, metabolic needs, and environmental conditions.

1.2 CHLOROPLAST PROTEOSTASIS

Protein homeostasis is a sum of protein synthesis, protein folding, and protein degradation. In this section, I will focus on protein degradation (proteolysis) and how it contributes to chloroplast proteostasis.

1.2.1 Plastids and proteolysis

Proteolysis impacts plastids in different ways (Figure 2A-B). First, proteolysis helps protein maturation within plastids. The N-terminal processing peptidase in stroma (SPP) and thylakoid lumen (TPP) cleave the signal peptides from the precursor proteins on arrival to their destination. The N-terminal aminopeptidase MAP1/2 modify the N-terminal methionine of the plastid-encoded proteins and stabilize proteins within the plastid. The C-terminal processing peptidase CtpA processes plastid-encoded D1 protein for its ligand binding and maturation. Second, proteolysis removes unwanted proteins or peptides within plastids. Organellar oligo-peptidase (Oop) and Presequence protease (Prep) process the cleaved transit peptides in stroma and avoid accumulation of the peptides that may be toxic to the cells. Third, proteolysis controls protein quality in plastids at various growth stage and environmental conditions. Central in proteostasis is the ATP-dependent CLP protease system, involved in *e.g.* removing misfolded proteins such as deoxyxylulose 5-phosphate synthase (DXS). In addition, FtsH and Deg are involved in turnover of damaged D1 protein of the photosystem reaction center. Fourth, proteolysis may release membrane-bound proteins (such as transcriptional regulators). Thylakoidal protease EGYs belong to M50A family. This family is characterized by its member *E. coli* RseP, which cleaves the intramembrane protein RseA and triggers cell signaling.

It is likely EGYs play a similar role. Fifth, extra-plastid proteolysis through autophagy or chlorophagy degrades the whole plastid or selected plastid in the vacuole, which recycles and relocates the nutrients from plastid to the cell (*14, 21*).

1.2.1.1 The chloroplast protease network

In Arabidopsis chloroplasts, at least three proteases have been identified in each compartment (envelope, stroma, thylakoid membrane, lumen); however, very little is known for most of these proteases except for their functional domains (through sequence alignment with known homologs) and their subcellular compartment (experimental data or predicted through co-expression network) (*22*) (Figure 2B). It is likely that several chloroplast proteases may share the same protein substrates but work sequentially or in parallel to process proteins. One example is the Deg and FtsH proteases cooperating in D1 protein turnover. D1 protein of the PSII reaction center is an integral thylakoid membrane protein. It has five transmembrane helixes (A-E) connected by loops exposed to either the stroma or lumen. It has been shown that specific D1 fragments accumulated in *var2* (*ftsh2*) mutant but not in the *deg2 deg8* mutants. Since thylakoidal FtsH1/2/5/8 form a FtsH complex degrading D1, and Deg2/8 locate to the lumen where the D1 loops and C-terminal are located, it was proposed that processing of the D1 by Deg can assist D1 degradation by FtsH (*23, 24*). Another possible network is CLP, OOP, and PREP, based mostly on their stromal localization, genetic interaction, and their substrate size. X-ray structures of PREP/OPP/CLP homologs suggests that PREP and OOP can degrade about 8-23 amino peptide substrates, whereas CLP can degrade bigger protein substrates and release peptide products around 8-12 amino acid (*25-28*). It has been proposed that

OOPs and PREPs function downstream of CLP. Whether these aminopeptidases work in parallel or sequentially is unknown (*14*).

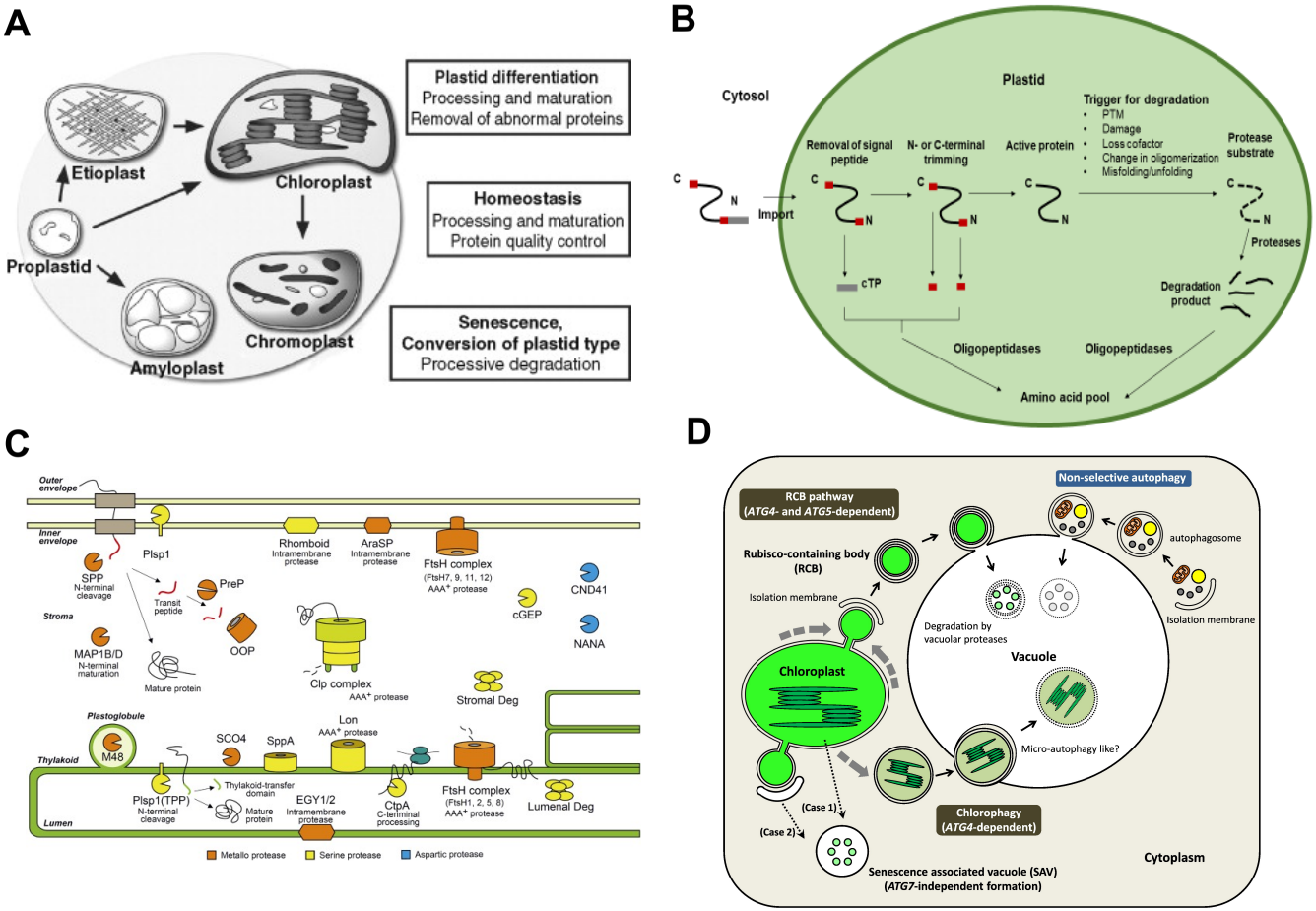
1.2.1.2 Chlorophagy and ATG-independent vesicles

In contrast to the intra-plastid proteolysis, extra-plastid degradation also regulates plastid protein homeostasis by degrading either intact chloroplasts or selected content of the chloroplasts in the vacuole (Figure 2D). This extra-plastid degradation is also known as autophagy and has been well characterized in yeast containing more than 30 autophagy-related genes (ATGs) involving in the process. During starvation, part of the yeast cytosolic content or organelles in the cytosol is engulfed by membrane. This forms a vesicle with double layer membranes (autophagesome). This autophagesome is then delivered to the vacuole for macromolecule degradation. In Arabidopsis, almost all yeast core ATG homologs are found in its genome (except for ATG14). Reverse genetics of these *atg* mutants and live imaging of the GFP-fusion ATG8 (autophagesome marker) showed that a similar autophagy pathway is present in plants, and this pathway is active in response to stresses such as nutrient starvation (*31*).

ATG-dependent pathways are involved in the regulation of chloroplast protein homeostasis. In the RCB (Rubisco-containing body) pathway and the chlorophagy pathway, partial content of the chloroplast stroma or the whole organelle is enclosed by autophagesomes and delivered to vacuoles. These two pathways need ATGs and can be found during senescence. Another type of senescence-induced pathway involves ATG-independent lytic vacuoles (senescence-associated vacuole; SAV) containing stromal content. SAV differs from RCB in its morphology and content

Figure 2. Proteases and plastids.

A, Proteases contribute to plastid life and death (*29*). Reprinted from International Review of Cell and Molecular Biology, Vol. 280, Yusuke Kato, Wataru Sakamoto, New insights into the types and function of proteases in plastids, p.187. Copyright © 2010, with permission from Elsevier. B, Protein maturation and proteolysis in chloroplasts. C, Chloroplast proteases (*30*). Reprinted by the permission from Plant Physiology, copyright by the American Society of Plant Biologists. D, Extra-plastid degradation pathways (*31*). Reprinted from Biochimica et Biophysica Acta, Vol. 1837, Hiroyuki Ishida, Masanori Izumi, Shinya Wade, Amane Makino, Roles of autophagy in chloroplast recycling, p.514. Copyright © 2014, with permission from Elsevier.


A
Plastid differentiation
Processing and maturation
Removal of abnormal proteins
Homeostasis
Processing and maturation
Protein quality control
Senescence,
Conversion of plastid type
Processive degradation
Etioplast
Chloroplast
Proplastid
Chromoplast
Amyloplast
B
Cytosol
Plastid
Trigger for degradation
PTM
Damage
Loss cofactor
Change in oligomerization
Misfolding/unfolding
Removal of signal peptide
N- or C-terminal trimming
Active protein
Protease substrate
Import
cTP
Proteases
Degradation product
Oligopeptidases
Oligopeptidases
Amino acid pool
C
Outer envelope
Inner envelope
Stroma
Thylakoid
Lumen
Plastoglobule
Plsp1
SPP
N-terminal cleavage
Transit peptide
PreP
OOP
MAP1B/D
N-terminal maturation
Mature protein
Rhomboid
Intramembrane protease
AraSP
Intramembrane protease
FtsH complex
(FtsH7, 9, 11, 12)
AAA+ protease
cGEP
CND41
NANA
Clp complex
AAA+ protease
Stromal Deg
M48
SCO4
SppA
Lon
AAA+ protease
Plsp1(TPP)
N-terminal cleavage
Thylakoid-transfer domain
Mature protein
EGY1/2
Intramembrane protease
CtpA
C-terminal processing
FtsH complex
(FtsH1, 2, 5, 8)
AAA+ protease
Lumenal Deg
Metallo protease
Serine protease
Aspartic protease
D
RCB pathway
(ATG4- and ATG5-dependent)
Non-selective autophagy
Rubisco-containing body
(RCB)
autophagosome
Isolation membrane
Isolation membrane
Degradation by vacuolar proteases
Vacuole
Chloroplast
Micro-autophagy like?
(Case 1)
(Case 2)
Chlorophagy
(ATG4-dependent)
Cytoplasm
Senescence associated vacuole (SAV)
(ATG7-independent formation)

SAV is enclosed by a monolayer membrane and carries a cysteine protease and chlorophyll a, which are not found in RCB. The mechanisms for these extra-plastid pathways are not known (*31-33*).

1.3. CHLOROPLAST CLP MACHINERY IN ARABIDOPSIS

1.3.1 The architecture of the chloroplast CLP machinery and its homologs

The CLP (caseinolytic protease) machinery is present in all prokaryotes and the organelles that originated from endosymbiosis, such as plastids and mitochondria, and has a conserved architecture across different species (*34-36*). This machinery is composed of several different components including adaptors, chaperones, and protease (*34, 35, 37*).

The CLP proteolytic core (protease) is a tetradecamer with two heptameric rings stacked together (*38, 39*). In bacteria and mitochondria, most of the CLP tetradecamer proteolytic cores are homo-oligomeric, composed of 14 identical CLPP monomers (*38, 40*). Each monomer can be divided into three parts: the N-terminal region, the globular head domain, and the handle region (*41*). X-ray structure of bacterial CLP showed that the main body of each heptameric ring is built by the head domain of 7 CLPP monomers and the interaction between the two heptameric rings is contributed by the dynamic handle region (*41*). The tetradecameric proteolytic CLPR core in the Arabidopsis chloroplast is a ~350 kDa complex, composed of 5 different CLPP (P1, P3-6) and 4 different CLPR (R1-4) subunits (*42*). These CLPP/R subunits form two heptameric rings: the P-ring (P3: P4: P5: P6= 1: 2: 3: 1) and the R-ring (P1: R1: R2: R3: R4= 3: 1: 1: 1) (*43*) (Figure 3A). It is unclear why the chloroplast Clp

protease has such a complicated composition and how each CLPP/R subunit is arranged in the 3D structure. Compared to the prokaryotic CLPP monomer, chloroplast CLPP/R subunits have a unique C-terminal extension except for the plastid-encoded CLPP1 (*40*). This C-terminal extension is common in higher plants, but the function of the C-terminal extension is unknown (*43*). Two accessory proteins named CLPT1/T2, which are homologous to the N-terminal region of AAA+ chaperones, are also unique to higher plants and are strongly associated with the CLPPR tetradecamer (*42, 44*). This association is likely due to the MYFF motif in CLPTs interacting with the hydrophobic cleft on CLPPR complexes (*44*).

CLP chaperones, as protease gatekeepers, self-associate into a hexamer and dock onto the CLP protease (*35*). In chloroplasts, AAA+ chaperones include CLPB3/C1/C2/D (*34*). All of these chaperones carry two ATPase domains and one IGF loop, except for CLPB, which lacks an IGF loop (also known as CLPP interacting motif) (*34*). It is believed that CLPC1/C2/D are involved in gating the CLPPRT complexes while CLPB3 plays another quality control role in removing the aggregated proteins by refolding. Interestingly, an R motif (required for cyanobacteria CLPC docking on the CLPP3/R complexes in cyanobacteria) is only found in CLPC1/C2 but not in CLPD, implying that CLPC1/C2 hexamer but not CLPD may interact with the chloroplast R-ring that originated from the cyanobacteria CLPP3/R complex (*45*). Adaptor proteins, known as scaffold or activator proteins, prepare substrates and protease for degradation (*46*). In Arabidopsis chloroplasts, two adaptor proteins, CLPS1 and CLPF, have been identified (*47, 48*). The core domain of CLPS1, the N-terminal domain and the UVR domain of CLPF are important for CLPC1 interaction

(*47*). How this bipartite adaptor system recognizes and delivers substrates is not fully understood.

1.3.1.1 Identification and evolution of the chloroplast CLPs

The whole chloroplast CLP protease complex was first identified as a 350 kDa complex in Arabidopsis and later in the non-photosynthetic plastid in Brassica (*40, 42*). This complex is composed of 10 different protein subunits named CLPP1/3-6, and CLPR1-4 according to the presence or absence of the catalytic triad (Ser-His-Asp)(*34*). Two additional CLPT proteins were discovered at the same time with the CLPPR complexes, but are considered to be outside the tetradecamer due to the low homology between CLPT and CLP/R proteins (*42*).

Sequence alignment demonstrated that the Arabidopsis CLPP/R subunits are not very conserved, sharing 20-40% identity within the CLPP or CLPR group (*49*). Phylogenic analysis suggested that most of these chloroplast CLPP/R originated from the cyanobacterial CLP1/P2 or CLPP3/R complexes (*34, 43*). Plastid-encoded CLPP1 and nuclear-encoded CLPR2 most likely evolved from the cyanobacteria CLPP3. Nuclear-encoded subunits CLPR1, 3, 4 are most similar to the cyanobacteria non-proteolytic CLPR. The origin of the nuclear-encoded CLPP6, as suggested by phylogenic analysis, is cyanobacterial CLPP1/P2. However, the origin of the nuclear-encoded CLPP3-5 and CLPT1-2 are still unclear. CLPT1/T2 are only found in higher plants (*34*). Chloroplast CLP chaperones (CLPC1/C2/D) and adaptors (CLPS1) were identified as bacterial homologs through sequence homology (*48, 50-52*). Arabidopsis CLPC1/2 are very similar to each other (88% identity) and to its cyanobacteria CLPC homolog (more than 70% identity); CLPD has ~ 45% sequence identity to the

cyanobacteria and chloroplast CLPC homologs (*34*). Adaptor CLPS1, detected by proteome analysis of Arabidopsis stroma, likely evolved from cyanobacteria CLPS1 (*48*). CLPF, on the other hand, was first identified as a CLPS1 putative substrate but later as an interactor since its interaction with CLPS1 does not require the substrate binding site of CLPS1 (*47*). CLPF is suggested as an evolutionary adaptation since it is found in all photosynthetic eukaryotes (*47*).

1.3.1.2 Reverse genetics of clp mutants

Homozygous *clpp/r* mutants display different levels of growth and molecular phenotypes, in part correlated with their stoichiometry in the tetradecamer (Figure 3B) (*49*). Losing *AtCLPP4* or *AtCLPP5* (two or three copies in the tetradecamer) caused an embryo lethal phenotype and an arrest at the globular stage during embryogenesis. Loss of expression of *AtCLPP3*, *AtCLPR2*, or *AtCLPR4* (one copy each) led to a seedling lethal phenotype. These homozygous mutants stopped growing at the cotyledon stage but could be partially rescued by a supplement of sucrose, suggesting that the chloroplast biogenesis was impaired in these mutants. *clpr1* nulls grew autotrophically but had a delayed growth and yellow-green rosette phenotype. The relatively moderate phenotype of *clpr1* is likely due to the redundancy between *CLPR1* and *CLPR3*. Complementation tests showed that *clpr1* could be recovered by overexpressing *CLPR3 in vivo*, but this is not the case for the other CLPP/R subunits. Since the other CLP/R subunits could not substitute for each other, these CLPP/R subunits are not redundant and may have a specific contribution toward the CLPPRT

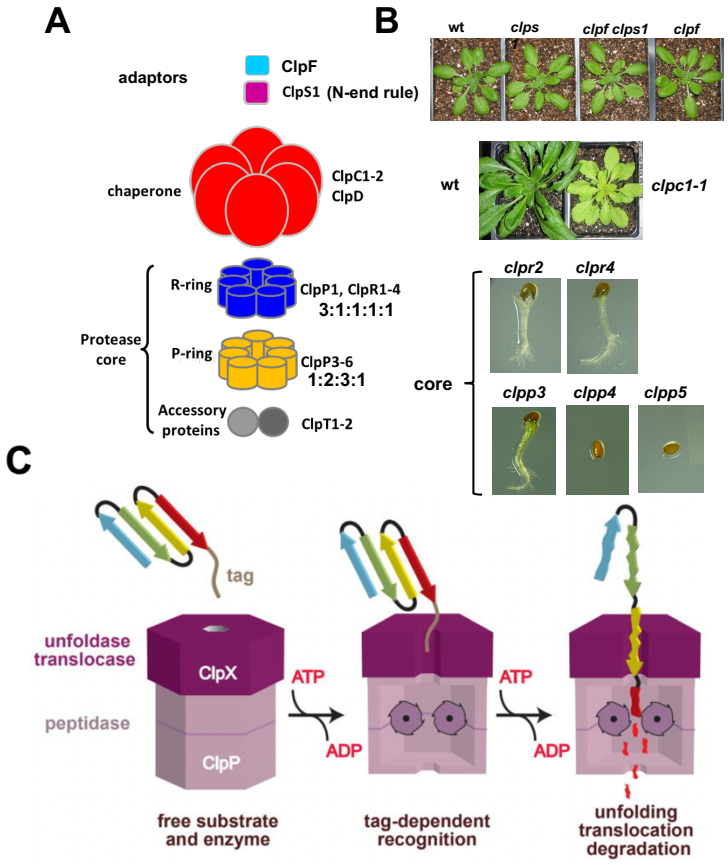


Figure 3. The architecture, components, and current model of CLP proteases.

A, The architecture of chloroplast CLP machinery in Arabidopsis (*34*). Reprinted from Biochimica et Biophysica Acta, Vol. 1847, Kenji Nishimura and Klaas J. van Wijk, Organization, function and substrates of the essential CLP protease system in plastids, p.916. Copyright © 2015, with permission from Elsevier. B, Phenotypes of the Arabidopsis *clp* nulls (*48, 49*). Reprinted by the permission from Plant Cell, copyright by the American Society of Plant Biologists. C, Substrate recognition and degradation mechanism of CLP machinery (*35*). Reprinted from Biochimica et Biophysica Acta, Vol. 1823, Tania A. Baker and Rebert T. Sauer, ClpXP, an ATP-powered unfolding and protein-degradation machine, p.16. Copyright © 2012, with permission from Elsevier.

machinery. Since there is no *CLPP6* T-DNA line and due to the difficulty of plastid gene transformation in Arabidopsis, the possible function of CLPP6 and CLPP1 was was deduced from the phenotypic analysis of tobacco transgenic lines or Arabidopsis antisense lines. Knock down of *CLPP6* in Arabidopsis and tobacco resulted in pale leaves with a retarded growth rate (*53, 54*). Downregulation of the plastid-encoded CLPP1 in tobacco resulted in a delayed growth and pigment deficiencies (*54*). Deletion of the whole *CLPP1* in tobacco led to a lethal phenotype (*55*). The reverse genetics data indicates that a) all the CLPP/R subunits are important for plant growth and development and b) each of them may have a unique contribution to the tetradecamer.

In contrast to the severe phenotypes observed in the tetradecamer mutants, the regulator (chaperones, adaptors, or accessory proteins) mutants are viable. Knocking out CLP chaperone (CLPD) or adaptors (CLPS1 or CLPF) does not result in a visible growth phenotype (*47, 48*). Only *clpc1* null but not *clpc2* displayed a pale green rosette phenotype, implying that CLPC1 plays a more important role in protein regulation (*48, 56*). Although no visible phenotype was observed in *clpc2* nulls, the dosage effect of *CLPC2* on *clpc1* indicated their genetic interaction *in vivo* (*48*). A synergistic effect between *clpt1clpt2* has also been observed. The *clpt1* or *clpt2* single mutant is WT-like, but *clpt1clpt2* nulls showed a virescent and serrated leaves with reduced growth (*44*).

1.3.2 Substrate recognition and degradation mechanisms of CLP systems

1.3.2.1 Substrate recognition and delivery

Substrate recognition and delivery for proteolytic degradation (Figure 3C) must be carried out timely and effectively in response to various environmental and developmental cues. CLP chaperones and adaptors play a major role in substrate recognition and delivery to the CLP proteolytic cores. CLP chaperones (as gate-keepers of the proteolytic chamber) selectively recognize the degradation signal(s) in substrates through its specified N-terminal substrate-interacting region and deliver the substrates to the proteolytic chamber (*35*). Two examples for this type of recognition are EcCLPA recognizing the N-terminus of RepA, and EcCLPX recognizing the C-terminal motif of the bacteriophage Mu repressor (*57, 58*). Adaptors (as scaffolds or activators) recognize degradation signals on the substrates and deliver substrates to protease for degradation (function as scaffolds) and/or prime whole CLP machinery for degradation (function as activators) (*46*). EcSspB is a scaffold adaptor that targets substrates with an ssrA-tag generated by the tmRNA when protein translation is incomplete (*59*). EcClpS is another example of scaffold adaptors. EcClpS recognizes the N-end rule degrons of the substrates (*60, 61*). BsMecA, on the other hand, acts as a priming adaptor that not only recognizes the substrates (ComK and ComS) but also activates the CLPCs and CLPPs to form a complex for substrate degradation (*62*).

In chloroplast, three CLP chaperones (AtCLPC1/C2/D) are believed to be involved in substrate recognition and delivery. However, the degradation signals recognized by each chaperone and the mechanism for substrate delivery to AtCLPPRT by each chaperone remain unclear. A recent X-ray structure showed that the N-terminal domain (substrate-interacting domain) of AtCLPD is structurally divergent from AtCLPC1, implying that these chaperones may have different substrate

recognition mechanisms (*63*). AtCLPD does not have an R-motif that is necessary for the hexamer to dock onto the cyanobacterial CLPP3R ring (the ancestor of the CLP R-ring), implying that CLPC1 and CLPD interact with the CLPPRT complex differently. It is possible that CLPD forms mixed oligomers with CLPC1,2 thus allowing it to participate in delivery of substrates. The different temporal presence of the CLPC and CLPD *in vivo* also hints that CLPD may play a unique role in recognizing the senescence-induced or stress-induced protein substrates (*52*).

Adaptor protein AtCLPS1, a bacterial CLPS homolog that recognizes and delivers the N-end rule substrates to bacterial CLPAP system, has been proposed to form a bipartite system with CLPF for degradation of glutamyl-tRNA reductase (GluTR) (*47*). The interacting domains among chaperone (CLPC1), adaptors (CLPS1 and CLPF), and the substrate (GluTR) have been identified through intensive biochemical studies *in vitro* using recombinant proteins. However, the degrons and the substrate specificity of these adaptors are still unclear. Whether a similar N-end rule presents in chloroplasts is unknown. Recently, a chloroplast N-terminome analysis and an *in vitro* CLPS1 affinity analysis suggest that several residues (F,L,W) may function as the N-degron for the CLPPRT system in chloroplasts (*64*) (Montandon et al, unpublished). Investigation of the recognition mechanism of degrons will advance our understanding of the protein homeostasis regulated by CLP protease in chloroplasts.

1.3.2.2 Substrate unfolding and translocation through chaperone gates

Chaperones control substrate entry into the proteolytic CLP chamber in two ways. First, chaperone hexamers can unfold and translocate substrates in an ATP-

dependent way (*35*). This is crucial for the folded protein substrates to enter the proteolytic chambers through the protease axial pores (~10 angstrom) (*65*). CLP chaperones have ATPase domain(s); each ATPase domain has two conserved motifs (walker A and walker B) involved in ATP binding and hydrolysis that empower the chaperones as the molecular motors to unfold the substrates and translocate the substrates into the CLP proteases (*35*). Second, chaperone hexamer can enlarge the axial pores of the protease when it docks onto the protease (*66*). The assembly of the chaperone hexamer and the CLP tetradecamer needs ATP binding but does not need ATP hydrolysis (*35*). This assembly, sometime aided by an adaptor (e.g. MecA that primes BsCLPCP), can widen the axial pores of the protease and activate the chaperones and protease for the subsequent degradation (*66-68*). Several motifs or regions involved in the interaction between the CLP chaperones and the protease have been identified for EcCLPXP. These include the IGF motif and the pore 2 loop of CLPX, as well as the hydrophobic cleft and the N-terminal loop of CLPP. The IGF loop can interact with the hydrophobic cleft of the CLPP complex, whereas the pore 2 loop can interact with the N-terminal loop of CLPP (*35, 69, 70*). It is unclear how chloroplast CLPC1/C2/D forms the hexamer (homo-oligo or hetero-oligomer) and where these hexamers dock onto the CLPPRT core. Whether other motifs or components stabilize the chloroplast CLPC-CLPPRT interaction is also unknown.

1.3.2.3 Substrate processing within the proteolytic barrel

The proteolytic barrel has two main structural characters that control the protein processing before and after entering into the proteolytic barrel. One is its narrow substrate entry pores (axial pores) contributed by the N-terminal loop of

CLPP; the other is its sequestered catalytic sites within the barrel-like tetradecamer (*65*). Therefore, only translocated substrates or small peptides can access the CLPP catalytic sites in the barrel. These two characters also prevent the degradation of the surrounding proteins of CLP proteases.

The CLPP catalytic site is a serine type catalytic triad (Ser-His-Asp) that is conserved across species and has also been found in other serine type proteases such as trypsin and chymotrypsin (*71*). Peptide bond cleavage by the CLP protease does not require ATP and is believed to be nonspecific cleavage because of a) the symmetrically homo-oligomeric composition of CLPP tetradecamer and b) the high density of the catalytic sites within the tetradecamer (14 catalytic sites in a spherical chamber with a 50 angstrom diameter) (*41, 65*). Interestingly, *Listeria monocytogenes* has two CLPP isoforms and can form homo-oligomer CLPP2 or heterooligomer CLPP1/P2 complexes *in vitro* (*72*). Peptide library screening using LmCLPP complexes showed a peptide cleavage preference between CLPP1 and CLPP2 but this preference becomes less distinct in the protease degradation assay (*73*). Whether the different CLPP isoforms in the same hetero-oligomeric CLP complex contribute to catalysis differently is not understood.

Once substrates enter the tetradecamer, they become accessible to the catalytic sites in the proteolytic chamber. The product of the CLP protease is around 8-12 amino acids. This size is predicted by the distance of the catalytic sites of the EcCLPP14 X-ray structure (25 angstrom). The product is most likely to exit through lateral sites between the two heptameric rings since the top and bottom ring surfaces are the docking sites of the CLP chaperones. X-ray structure of the BsCLPP and

SaCLPP showed several conformational states of CLPP complexes, which likely reflect the dynamic conformational states of CLPP complexes *in vivo*. In the extended state, the two heptameric rings of the CLP complex were connected by the handle regions through the hydrogen bond network and the catalytic triads of the CLPPs were oriented properly. It is therefore believed that CLP protease in this conformational state is active. On the contrary, BsCLPP and SaCLPP in the compressed stated lost the connection between the handle regions seen in the extend state. The side pores between the two rings appeared instead. Because the catalytic triads of the BsCLPP and SaCLPP in this state were all miss-oriented, this conformational state is believed as an inactive state. The third state is the compact state. This state has been proposed as the intermediate state, a transition from the extend state to the compressed state. CLPP complexes in this state did not have their catalytic triads oriented properly and their handle regions (E-helix) did not break as seen in the compressed form. These structural studies suggest that CLPP *in vivo* may undergo a dynamic conformational change and the protease product is possibly released from the lateral pore openings in the compressed conformational state (*41*).

Although the intensive structural and biochemical studies have been done on bacterial CLPPs, little is known about the protein 3D structure and degradation mechanism of the chloroplast CLPPRT except for their stoichiometry. The main challenges to obtain structural information is the complicated composition of the chloroplast CLPPRT and low yield of the CLPPRT purified from plants for subsequent structural and kinetics analysis. Alternative approaches using structural

proximity crosslinking, and co-evolution analysis of protein subunits within the CLPPRT complexes may help to resolve the 3D structure of the CLPPRT complexes.

1.3.3 Known CLP substrates and the approaches used for substrate discovery

So far, only a few chloroplast CLP substrates have been identified. GluTR (Glutamyl-tRNA reductase) was first identified as a direct interactor using *in vitro* trapping, combining affinity-MS analysis of the wt CLPS1 and mutated ClpS1 carrying point mutations in the CLPS1 substrate binding (*48*). DXS (deoxyxylulose 5-phosphate synthase) and PAAM2/HMA8 (P-type ATPase of Arabidopsis 2 or Heavy Metal ATPase 8) were discovered through their overaccumulation in *clp* mutants (*74, 75*). PSY (phytoene synthase) was discovered as an interactor with CLPC through pull down-MS (*76*). These candidates were all further tested for their subcellular location, protein-protein interaction with the CLP system, and the protein turnover in *clp* mutants.

Substrate candidate lists have been also generated by comparative proteomic approaches (*77-79*). These approaches, in principle, investigate the proteomic difference between two cells (wt versus mutant) or two stages (inhibitor before added versus inhibitor after added). Putative substrates were identified as those proteins accumulated in the protease mutants or enriched after protease inhibitor was added. One benefit of this approach is that it provides the peptide information that may include the possible protease cleavage site. These *clp* mutant lines were either generated by knocking out a CLP subunit or temporally suppressing the CLP expression (*77, 79, 80*). However, this approach can not distinguish between primary and secondary effects caused by the impairment of the CLP machinery *in vivo*.

Recently, new AAA+ protease substrates have been identified using *in vivo* trapping (*81-83*). The rationale of the *in vivo* trapping is to prevent substrates from degradation and enrich the substrates trapped by the protease machinery for affinity-MS analysis. With an in parallel control, putative substrates are those proteins that are only present or highly enriched in the trap sample but not in the control line. This method keeps the advantages of the proteomic approaches and overcomes the secondary effect of the comparative proteomic analysis of the total cell extracts. CLPPRT is an abundant protease in the chloroplast stroma and is constitutively expressed. Systematic identification of the chloroplast CLP substrates and their degradation signals are key to understand and control the CLP machinery in chloroplasts.

1.4 THESIS OBJECTIVES AND AIMS OF THE CHAPTERS

The objective of this dissertation is to investigate the function and structure of the chloroplast CLP machinery through substrate identification and investigation of the spatial arrangement of CLP machinery. Systematic substrate identification provides an unbiased way to unravel protease biological functions *in vivo*. Chapter two reviews the experimental studies of protease substrate trapping using an *in vivo* trapping approach. The aim of chapter two is to summarize the key features and limitations of *in vivo* trapping, as well as to generate a guideline for substrate identification using this approach. Chapter three investigates the chloroplast household CLP protease using inactive CLPP3 and CLPP5 subunits for *in vivo* trapping. The aim of this chapter is to understand the functional contribution of CLPP subunits and identify substrate candidates of the chloroplast CLP. Chapter four studies the

proximity of the CLPPRT and CLPC using DSSO-MSn analysis. This chapter aims to provide preliminary data for further investigation of the spatial arrangement of the individual subunits in the tetradecamer using co-evolution analysis. Chapter five summarizes all chapters and provides future directions of chloroplast CLP research from a structural and functional perspective.

REFERENCES

1. T. Cavalier-Smith, Only six kingdoms of life. *Proceedings of the Royal Society B-Biological Sciences* **271**, 1251-1262 (2004).
2. A. Reyes-Prieto, A. P. M. Weber, D. Bhattacharya, The origin and establishment of the plastid in algae and plants. *Annual Review of Genetics* **41**, 147-168 (2007).
3. P. Jarvis, E. Lopez-Juez, Biogenesis and homeostasis of chloroplasts and other plastids. *Nature Reviews: Molecular Cell Biology* **14**, 787-802 (2013).
4. L. Li, H. Yuan, Chromoplast biogenesis and carotenoid accumulation. *Archives of Biochemistry and Biophysics* **539**, 102-109 (2013).
5. J. P. Carde, Leucoplasts: a distinct kind of organelles lacking typical 70S ribosomes and free thylakoids. *European Journal of Cell Biology* **34**, 18-26 (1984).
6. R. Chen, C. Guan, K. Boonsirichai, P. H. Masson, Complex physiological and molecular processes underlying root gravitropism. *Plant Molecular Biology* **49**, 305-317 (2002).
7. L. A. Staehelin, Chloroplast structure: from chlorophyll granules to supra-molecular architecture of thylakoid membranes. *Photosynthesis Research* **76**, 185-196 (2003).
8. B. J. Pogson, D. Ganguly, V. Albrecht-Borth, Insights into chloroplast biogenesis and development. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **1847**, 1017-1024 (2015).
9. K. Inoue, Emerging roles of the chloroplast outer envelope membrane. *Trends in Plant Science* **16**, 550-557 (2011).
10. U. C. Vothknecht, P. Westhoff, Biogenesis and origin of thylakoid membranes. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **1541**, 91-101 (2001).
11. K. J. van Wijk, F. Kessler, Plastoglobuli: Plastid microcompartments with integrated functions in metabolism, plastid developmental transitions, and environmental adaptation. *Annual Review of Plant Biology* **68**, 253-289 (2017).

12. U. Flores-Perez, P. Jarvis, Molecular chaperone involvement in chloroplast protein import. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **1833**, 332-340 (2013).
13. G. Lamberti, C. Drurey, J. Soll, S. Schwenkert, The phosphorylation state of chloroplast transit peptides regulates preprotein import. *Plant Signal Behav* **6**, 1918-1920 (2011).
14. K. J. van Wijk, Protein maturation and proteolysis in plant plastids, mitochondria, and peroxisomes. *Annual Review of Plant Biology* **66**, 75-111 (2015).
15. A. Barkan, Expression of plastid genes: organelle-specific elaborations on a prokaryotic scaffold. *Plant Physiology* **155**, 1520-1532 (2011).
16. M. R. Groves *et al.*, Functional characterization of recombinant chloroplast signal recognition particle. *Journal of Biological Chemistry* **16**, 16 (2001).
17. J. W. Bradbeer, Y. E. Atkinson, T. Borner, R. Hagemann, Cytoplasmic synthesis of plastid polypeptides may be controlled by plastid-synthesised RNA. *Nature* **279**, 816 (1979).
18. L. Tadini *et al.*, GUN1 controls accumulation of the plastid ribosomal protein s1 at the protein level and interacts with proteins involved in plastid protein homeostasis. *Plant Physiology* **170**, 1817-1830 (2016).
19. A. de Souza, J. Z. Wang, K. Dehesh, Retrograde signals: integrators of interorganellar communication and orchestrators of plant development. *Annual Review of Plant Biology* **68**, 85-108 (2017).
20. W. Chi, X. W. Sun, L. X. Zhang, Intracellular signaling from plastid to nucleus. *Annual Review of Plant Biology, Vol 64* **64**, 559-582 (2013).
21. K. Nishimura, Y. Kato, W. Sakamoto, Essentials of proteolytic machineries in chloroplasts. *Mol Plant* **10**, 4-19 (2017).
22. K. Majsec *et al.*, The plastid and mitochondrial peptidase network in Arabidopsis thaliana: a foundation for testing genetic interactions and functions in organellar proteostasis. *Plant Cell* **29**, 2687-2710 (2017).

23. Y. Kato, X. Sun, L. Zhang, W. Sakamoto, Cooperative D1 degradation in the photosystem II repair mediated by chloroplastic proteases in Arabidopsis. *Plant Physiology* **159**, 1428-1439 (2012).
24. Y. Kato, E. Miura, K. Ido, K. Ifuku, W. Sakamoto, The variegated mutants lacking chloroplastic FtsHs are defective in D1 degradation and accumulate reactive oxygen species. *Plant Physiology* **151**, 1790-1801 (2009).
25. B. Kmiec *et al.*, Organellar oligopeptidase (OOP) provides a complementary pathway for targeting peptide degradation in mitochondria and chloroplasts. *Proceedings of the National Academy of Sciences of the United States of America* **110**, E3761-3769 (2013).
26. P. F. Teixeira, E. Glaser, Processing peptidases in mitochondria and chloroplasts. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **1833**, 360-370 (2013).
27. B. Kmiec, P. F. Teixeira, E. Glaser, Shredding the signal: targeting peptide degradation in mitochondria and chloroplasts. *Trends in Plant Science* **19**, 771-778 (2014).
28. E. Rowland, Dissertation, Cornell University, (2017).
29. Y. Kato, W. Sakamoto, New insights into the types and function of proteases in plastids. *International Review of Cell and Molecular Biology* **280**, 185-218 (2010).
30. K. Nishimura, Y. Kato, W. Sakamoto, Chloroplast Proteases: updates on proteolysis within and across suborganellar compartments. *Plant Physiology* **171**, 2280-2293 (2016).
31. H. Ishida, M. Izumi, S. Wada, A. Makino, Roles of autophagy in chloroplast recycling. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **1837**, 512-521 (2014).
32. H. Ishida, S. Wada, Autophagy of whole and partial chloroplasts in individually darkened leaves: a unique system in plants? *Autophagy* **5**, 736-737 (2009).

33. H. Ishida, K. Yoshimoto, Chloroplasts are partially mobilized to the vacuole by autophagy. *Autophagy* **4**, 961-962 (2008).
34. K. Nishimura, K. J. van Wijk, Organization, function and substrates of the essential Clp protease system in plastids. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **1847**, 915-930 (2015).
35. T. A. Baker, R. T. Sauer, ClpXP, an ATP-powered unfolding and protein-degradation machine. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **1823**, 15-28 (2012).
36. S. L. Wang, X. Q. Liu, Identification of an unusual intein in chloroplast ClpP protease of Chlamydomonas eugametos. *Journal of Biological Chemistry* **272**, 11869-11873 (1997).
37. P. Chien, B. S. Perchuk, M. T. Laub, R. T. Sauer, T. A. Baker, Direct and adaptor-mediated substrate recognition by an essential AAA+ protease. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 6590-6595 (2007).
38. J. Wang, J. A. Hartling, J. M. Flanagan, Crystal structure determination of Escherichia coli ClpP starting from an EM-derived mask. *Journal of Structural Biology* **124**, 151-163 (1998).
39. A. Gribun *et al.*, The ClpP double ring tetradecameric protease exhibits plastic ring-ring interactions, and the N termini of its subunits form flexible loops that are essential for ClpXP and ClpAP complex formation. *Journal of Biological Chemistry* **280**, 16185-16196 (2005).
40. J. B. Peltier *et al.*, Clp protease complexes from photosynthetic and non-photosynthetic plastids and mitochondria of plants, their predicted three-dimensional structures, and functional implications. *Journal of Biological Chemistry* **279**, 4768-4781 (2004).
41. K. Liu, A. Ologbenla, W. A. Houry, Dynamics of the ClpP serine protease: a model for self-compartmentalized proteases. *Critical Reviews in Biochemistry and Molecular Biology* **49**, 400-412 (2014).

42. J. B. Peltier, J. Ytterberg, D. A. Liberles, P. Roepstorff, K. J. van Wijk, Identification of a 350-kDa ClpP protease complex with 10 different Clp isoforms in chloroplasts of Arabidopsis thaliana. *Journal of Biological Chemistry* **276**, 16318-16327. (2001).

43. P. D. Olinares, J. Kim, J. I. Davis, K. J. van Wijk, Subunit stoichiometry, evolution, and functional implications of an asymmetric plant plastid ClpP/R protease complex in Arabidopsis. *Plant Cell* **23**, 2348-2361 (2011).

44. J. Kim *et al.*, Structures, functions, and interactions of ClpT1 and ClpT2 in the Clp protease system of Arabidopsis chloroplasts. *Plant Cell* **27**, 1477-1496 (2015).

45. A. Tryggvesson, F. M. Stahlberg, A. Mogk, K. Zeth, A. K. Clarke, Interaction specificity between the chaperone and proteolytic components of the cyanobacterial Clp protease. *Biochemical Journal* **446**, 311-320 (2012).

46. N. J. Kuhlmann, P. Chient, Selective adaptor dependent protein degradation in bacteria. *Current Opinion in Microbiology* **36**, 118-127 (2017).

47. K. Nishimura *et al.*, Discovery of a unique clp component, ClpF, in chloroplasts: a proposed binary ClpF-ClpS1 adaptor complex functions in substrate recognition and delivery. *Plant Cell* **27**, 2677-2691 (2015).

48. K. Nishimura *et al.*, ClpS1 is a conserved substrate selector for the chloroplast Clp protease system in Arabidopsis. *Plant Cell* **25**, 2276-2301 (2013).

49. J. Kim *et al.*, Subunits of the plastid ClpPR protease complex have differential contributions to embryogenesis, plastid biogenesis, and plant development in Arabidopsis. *Plant Cell* **21**, 1669-1692 (2009).

50. J. Shanklin, N. D. DeWitt, J. M. Flanagan, The stroma of higher plant plastids contain ClpP and ClpC, functional homologs of Escherichia coli ClpP and ClpA: an archetypal two-component ATP-dependent protease. *Plant Cell* **7**, 1713-1722 (1995).

51. K. Nakabayashi, M. Ito, T. Kiyosue, K. Shinozaki, A. Watanabe, Identification of clp genes expressed in senescing Arabidopsis leaves. *Plant & Cell Physiology* **40**, 504-514 (1999).

52. K. Nakashima, T. Kiyosue, K. Yamaguchi-Shinozaki, K. Shinozaki, A nuclear gene, erd1, encoding a chloroplast-targeted Clp protease regulatory subunit homolog is not only induced by water stress but also developmentally up-regulated during senescence in Arabidopsis thaliana. *Plant Journal* **12**, 851-861 (1997).

53. L. L. Sjogren, T. M. Stanne, B. Zheng, S. Sutinen, A. K. Clarke, Structural and functional insights into the chloroplast ATP-dependent Clp protease in Arabidopsis. *Plant Cell* **18**, 2635-2649 (2006).

54. J. C. Moreno *et al.*, Generation and characterization of a collection of knock-down lines for the chloroplast Clp protease complex in tobacco. *Journal of Experimental Botany* **68**, 2199-2218 (2017).

55. H. Kuroda, P. Maliga, The plastid clpP1 protease gene is essential for plant development. *Nature* **425**, 86-89 (2003).

56. L. L. Sjogren, T. M. MacDonald, S. Sutinen, A. K. Clarke, Inactivation of the clpC1 gene encoding a chloroplast Hsp100 molecular chaperone causes growth retardation, leaf chlorosis, lower photosynthetic activity, and a specific reduction in photosystem content. *Plant Physiology* **136**, 4114-4126 (2004).

57. J. R. Hoskins, S. Y. Kim, S. Wickner, Substrate recognition by the ClpA chaperone component of ClpAP protease. *Journal of Biological Chemistry* **275**, 35361-35367 (2000).

58. J. E. Laachouch, L. Desmet, V. Geuskens, R. Grimaud, A. Toussaint, Bacteriophage Mu repressor as a target for the Escherichia coli ATP-dependent Clp protease. *EMBO Journal* **15**, 437-444 (1996).

59. D. A. Dougan, E. Weber-Ban, B. Bukau, Targeted delivery of an ssrA-tagged substrate by the adaptor protein SspB to its cognate AAA+ protein ClpX. *Molecular Cell* **12**, 373-380 (2003).

60. G. M. De Donatis, S. K. Singh, S. Viswanathan, M. R. Maurizi, A single ClpS monomer is sufficient to direct the activity of the ClpA hexamer. *Journal of Biological Chemistry* **285**, 8771-8781 (2010).

61. M. A. Humbard, S. Surkov, G. M. De Donatis, L. M. Jenkins, M. R. Maurizi, The N-degradome of Escherichia coli: limited proteolysis in vivo generates a large pool of proteins bearing n-degrons. *Journal of Biological Chemistry* **288**, 28913-28924 (2013).
62. Z. Mei *et al.*, Molecular determinants of MecA as a degradation tag for the ClpCP protease. *Journal of Biological Chemistry* **284**, 34366-34375 (2009).
63. C. Mohapatra, M. Kumar Jagdev, D. Vasudevan, Crystal structures reveal N-terminal domain of Arabidopsis thaliana ClpD to be highly divergent from that of ClpC1. *Scientific Reports* **7**, 44366 (2017).
64. E. Rowland, J. Kim, N. H. Bhuiyan, K. J. van Wijk, The Arabidopsis chloroplast stromal n-terminome: complexities of amino-terminal protein maturation and stability. *Plant Physiology* **169**, 1881-1896 (2015).
65. J. Wang, J. A. Hartling, J. M. Flanagan, The structure of ClpP at 2.3 A resolution suggests a model for ATP-dependent proteolysis. *Cell* **91**, 447-456 (1997).
66. G. Effantin, M. R. Maurizi, A. C. Steven, Binding of the ClpA unfoldase opens the axial gate of ClpP peptidase. *Journal of Biological Chemistry* **285**, 14834-14840 (2010).
67. K. R. Schmitz, R. T. Sauer, Substrate delivery by the AAA+ ClpX and ClpC1 unfoldases activates the mycobacterial ClpP1P2 peptidase. *Molecular Microbiology* **93**, 617-628 (2014).
68. M. Carroni *et al.*, Regulatory coiled-coil domains promote head-to-head assemblies of AAA+ chaperones essential for tunable activity control. *Elife* **6**, (2017).
69. S. A. Joshi, G. L. Hersch, T. A. Baker, R. T. Sauer, Communication between ClpX and ClpP during substrate processing and degradation. *Nature Structural & Molecular Biology* **11**, 404-411 (2004).
70. A. Martin, T. A. Baker, R. T. Sauer, Pore loops of the AAA+ ClpX machine grip substrates to drive translocation and unfolding. *Nature Structural & Molecular Biology* **15**, 1147-1151 (2008).

71. M. R. Maurizi, W. P. Clark, S. H. Kim, S. Gottesman, Clp P represents a unique family of serine proteases. *Journal of Biological Chemistry* **265**, 12546-12552 (1990).

72. M. Dahmen, M. T. Vielberg, M. Groll, S. A. Sieber, Structure and mechanism of the caseinolytic protease ClpP1/2 heterocomplex from Listeria monocytogenes. *Angewandte Chemie-International Edition* **54**, 3598-3602 (2015).

73. D. Balogh *et al.*, Insights into ClpXP proteolysis: heterooligomerization and partial deactivation enhance chaperone affinity and substrate turnover in Listeria monocytogenes. *Chemical Science* **8**, 1592-1600 (2017).

74. P. Pulido *et al.*, Specific Hsp100 chaperones determine the fate of the first enzyme of the plastidial isoprenoid pathway for either refolding or degradation by the stromal Clp protease in Arabidopsis. *PLoS Genetics* **12**, (2016).

75. W. Tapken, J. Kim, K. Nishimura, K. J. van Wijk, M. Pilon, The Clp protease system is required for copper ion-dependent turnover of the PAA2/HMA8 copper transporter in chloroplasts. *New Phytologist* **205**, 511-517 (2015).

76. R. Welsch *et al.*, Clp protease and OR directly control the proteostasis of phytoene synthase, the crucial enzyme for carotenoid biosynthesis in Arabidopsis. *Molecular Plant* **11**, 149-162 (2018).

77. J. C. Moreno *et al.*, Temporal proteomics of inducible RNAi lines of Clp protease subunits identifies putative protease substrates. *Plant Physiology* **176**, 1485-1508 (2018).

78. B. Zybailov *et al.*, Large scale comparative proteomics of a chloroplast Clp protease mutant reveals folding stress, altered protein homeostasis, and feedback regulation of metabolism. *Molecular and Cellular Proteomics* **8**, 1789-1810 (2009).

79. T. M. Stanne, L. L. Sjogren, S. Koussevitzky, A. K. Clarke, Identification of new protein substrates for the chloroplast ATP-dependent Clp protease supports its constitutive role in Arabidopsis. *Biochemical Journal* **417**, 257-268 (2009).

80. J. Kim *et al.*, Modified Clp protease complex in the ClpP3 null mutant and consequences for chloroplast development and function in Arabidopsis. *Plant Physiology* **162**, 157-179 (2013).
81. J. Feng *et al.*, Trapping and proteomic identification of cellular substrates of the ClpP protease in Staphylococcus aureus. *Journal of Proteome Research* **12**, 547-558 (2013).
82. J. Arends, N. Thomanek, K. Kuhlmann, K. Marcus, F. Narberhaus, In vivo trapping of FtsH substrates by label-free quantitative proteomics. *Proteomics* **16**, 3161-3172 (2016).
83. J. Arends *et al.*, An integrated proteomic approach uncovers novel substrates and functions of the Lon protease in Escherichia coli. *Proteomics* **18**, (2018).

CHAPTER TWO

DISCOVERY OF AAA+ PROTEASE SUBSTRATES THROUGH TRAPPING APPROACHES[1]

2.1 ABSTRACT

Proteases play essential roles in cellular proteostasis. Mechanisms through which proteases recognize their substrates are often hard to predict and therefore require experimentation. *In vivo* trapping allows systematic identification of potential substrates of proteases, their adaptors and chaperones. This combines *in vivo* genetic modifications of proteolytic systems, stabilization of protease-substrate interactions, affinity enrichments of trapped substrates, and mass spectrometry-based identification. *In vitro* approaches, in which immobilized protease components are incubated with isolated cellular proteome, complement this *in vivo* approach. Both approaches can provide information about substrate recognition signals, degrons, as well as conditional effects. This review summarizes published trapping studies and their biological outcomes, and provides recommendations for substrate trapping of the processive AAA+ Clp, Lon and FtsH chaperone-proteolytic systems.

---

[1] Jui-Yun Rei Liao and Dr. Klaas J. van Wijk wrote the manuscript. This work has been published in Trend in Biochemical Sciences, 2019, in press.

2.2 PROTEOLYTIC SYSTEMS AND SUBSTRATE SELECTION, UNFOLDING AND DEGRADATION

Cellular proteomes maintain protein homeostasis (proteostasis) through control of protein synthesis, maturation and degradation (*1*). A wide range of proteases exist which are classified using a universal system as outlined in the MEROPS database at https://www.ebi.ac.uk/merops/ (*2*). The main challenge in protease research and understanding cellular proteostasis is determining how substrates are recognized by proteases and how the activity of proteases is regulated. This recognition typically involves modification of proteins resulting in the formation of a degron (see Glossary). Such degrons can be the unfolded domain of a protein, or can involve one or more post-translational modifications (PTMs), *e.g.* poly-ubiquitination by E3-ligases in eukaryotes (*3-5*) or addition of an N-terminal residue by amino acid transferases such as Leu/Phe-tRNA proteins transferases in case of N-end rule substrates in prokaryotes (*6*), or addition of the C-terminal ssrA tag (11 amino acids) (*7, 8*) through the action of ssrB (*9*). Generally, it is difficult to predict which proteins are degraded by which protease system. Furthermore, degradation of proteins frequently involves more than one protease, either acting in series or in parallel. Yet, such hierarchies are poorly understood. Thus, experimentation is needed to determine the substrates for proteases.

An important class of proteases are the ATP-dependent AAA+ proteases (see Glossary), in particular the Clp protease system, the Lon and the FtsH proteases. It should be noted that members of these families may have alternative names in different species (*e.g.* FtsH members AFG3L2 and YME1 in human and yeast, respectively). These proteolytic systems are found in eubacteria and eukaryotes, and

substrate selection does not involve ubiquitination. They have been studied in eubacteria such as *Escherichia coli* (*E. coli*), *Staphylococcus aureus* (*S. aureus*), and *Bacillus subtilis* (*B. subtilis*) as well as in mitochondria and plastids/chloroplasts in a range of eukaryotes (*10-13*). These AAA+ proteases form barrel-like proteolytic chambers and rely on either separate or integrated AAA+ unfoldase/chaperones to interact, unfold, and translocate the substrates into the proteolytic chamber. They are often aided by so-called adaptor proteins (see Glossary) many of which select specific substrate classes and deliver them to the chaperone-protease complex (*14*). The processive nature of these proteases and the lack of cleavage site motifs makes prediction and experimental recognition of their substrates challenging.

This review is focused on protease substrate trapping in AAA+ protease systems, which is an effective way to identify substrates for these ATP-dependent proteolytic systems. Table 1 provides a complete inventory of published substrate trapping studies for Clp, FtsH and Lon system in prokaryotes (*E. coli, B., S. aureus*, *Caulobacter crescentus – S. crescentus),* in chloroplasts of the plant species *Arabidopsis thaliana* (*A. thaliana*), and in mitochondria of the fungus *Podospora anserine* (*P. anserine*), *A. thaliana*, and embryonic fibroblasts of mice. Substrate trapping is also effective and popular for E3 ligases that ubiquitinate specific proteins for selective degradation by the proteasome (*15*), but this will not be discussed here.

Whereas substrate trapping has been used in multiple studies, as will be reviewed in this paper, it could be used to discover many more AAA+ protease substrates across a much larger number of species, environmental, developmental conditions and genetic backgrounds. Moreover, AAA+ proteases in eukaryotes are

often represented by (large) gene families and the functional diversity within these families is poorly understood. As an example, the *A. thaliana* genome encodes for 17 FtsH/FtsH-like proteins, eight Lon/Lon-like proteins distributed mostly across chloroplasts and mitochondria, whereas the mitochondrial and chloroplast Clp system includes at least 17 proteins. Substrate trapping can help to identify the specific functions and substrates of components within each of these diversified systems. This review therefore summarizes key technical features and biological outcomes of the published studies of *in vivo* trapping for Clp, Lon and FtsH chaperone-proteolytic systems (Table 1). We also introduce and review a complementary *in vitro* trapping approach, which so far has only been used for the adaptor ClpS in *E. coli* and chloroplasts of *A. thaliana* (Table 1). The review also provides suggestions and recommendations for future studies.

2.3 PROTEOMIC APPROACHES TO DISCOVER PROTEASE SUBSTRATES AND CLEAVAGE SPECIFICITIES

To systematically identify candidate protease substrates and study protease-substrate relationships *in vivo*, several proteomics approaches are available (*16*), namely comparative quantitative proteomics and peptidomics, as well as N-terminomics using N-terminal labeling techniques such as TAILS and COFRADIC (see Glossary) (*17-22*). These approaches typically analyze and compare the proteome and peptidome composition in (sub)cellular protein extracts in the presence or absence of the protease (activity) or during a time course, *e.g.* for plant metacaspases (*23*). Potential substrates are then identified by quantitative and/or qualitative comparison of

Table 1. Overview and summary of highlights and key information of published substrate trapping studies for the AAA+ proteolytic systems of Clp, FtsH and Lon families.

Bait	Key points and novelty	Technical aspects and other comments	affinity tags	Species and substrate pool	reference
Adaptor trap					
ClpS (*in vitro*)	About 20 ClpS targets. Test degron binding specificity with ClpS DD/AA mutant and also different dipeptides for elution. Discovery of LFTR as N-terminal modifier of substrate PATase. ClpS substrates also have a hydrophobic element 7-10 AA downstream of the N-degron. This is needed for ClpAP degradation. Dps was N-terminally cleaved prior to becoming a ClpS substrate by unknown protease(s). Extensive follow-up experiments confirm substrates.	Maldi-TOF for GluC digest. MSMS for tryptic digest. *In vitro* trapping using immobolized ClpS(wt) and ClpS mutant (negative control for N-degron recognition) with clpA- cell extract. Both 1D and 2D gel to visualize ClpS interactors. Extensive Edman degradation sequencing and MS/MS to determini the N-termini of substrates. Elution with dipeptides.	His6	*E. coli* (a)	Ninnis *et al* 2009
ClpS (*in vitro*)	12 ClpS candidate targets. Test degron binding specificity with ClpS DD/AA mutant and also different dipeptides for elution. PATase (YgjG) and Dps were most abundant interactors and therefore follow-up experiments were pursued for these two candidate substrates, confirming the strict ClpS dependence for degradation of N-terminally modified forms and lack of degradation by CLPXP. The ClpS interacting Dps starts at Leu6 and this truncated form (not functional as it does bind to DNA) is not a target of ClpXP. PATase also strictly requires ClpS and stability of PATase in the *aat* mutant (null for LFTR/Aat) background suggests N-terminal modification of Arg3.	*In vitro* trapping using immobolized ClpS(wt) and ClpS mutant (negative control for N-degron recognition) with ClpS- cell extract. 1D-SDS PAGE followed by in-gel digestion of 16 bands per lane and LC-MSMS. Elution with dipeptides.	GST	*E. coli* (a)	*Schmidt et al* 2009
ClpS (*in vitro*)	Over 100 proteins preferentially bound to ClpS as compared to ClpS-DD/AA and eluted with a peptide bearing an N-degron. 32 out of 37 determined N-terminal peptides had N-degrons. Most of the proteins were N-terminally truncated by endoproteases or exopeptidases, and many were likely further modified by LFTR/Aat. Suggest roles for the N-end rule in cell division, translation, transcription, and DNA replication and reveal widespread proteolytic processing of cellular proteins to generate N-end rule substrates.	MSMS and spectral counting and Maldi-TOF. N-terminal Edman degradation sequencing of 2D gel separated protein spots identified many N-termini. Proteins were eluted with peptide FKTA-NH2. Statistical G-test for significance levels of enriched proteins; Impute for missing values. The cellular proteome was from wild-type stationary phase cultures. It would be interesting to repeat these experiments with log phase cultures and compare the substrates and degrons.	His6	*E. coli* (a)	Humbard *et al* 2013
ClpS1 (*in vitro*)	About nine high confidence candidate ClpS1 substrates based on comparison of ClpS1 and ClpS1-DN/AA (and other controls). These include several enzymes in the shikimate and GluTR in the tetrapyrole pathway. No obvious N-degron was detected. Discovery of novel candidate co-adaptor ClpF of ClpS1, unique to higher plant chloroplasts. Most candidate substrates have inceased levels in various clp-loss-of-function mutants.	Use chloroplast stroma of wt and the double null mutant *clpc1 clps1* . Independent replicates and quantification by spectral counting. Thresholds for candidate substrates include repeat observations and fold-enrichment.	GST	*A. thaliana* (b)	Nishimura *et al* 2013
RcdA (*in vivo*)	In vivo capture with tagged RcdA identified known substrates TacA, as well as CC2323 and CC3144 (and additional proteins that were not further studied). CC22323 was identified previously as a ClpXP substrate based on a ClpP trapping. CpdR and RcdA were both required for CC3144 degradation in vivo, Similarly, CC2323 in vitro degradation by ClpXP was enhanced by addition of CpdR and RcdA. Taken together, these data reveal that the RcdA adaptor can bind and deliver a number of substrates in addition to TacA.	M2 epitope-tagged RcdA variants M2-RcdA and M2-RcdA-dC (lacking the C-terminal required for primed ClpX interaction), were expressed in *drcda* cells. After cell lysis, RcdA-interacting proteins were enriched using M2-FLAG affinity beads and proteins identified by 1D-SDS PAGE and MS/MS. TacA and candidate targets were (somewhat) enriched in the RcdA-dC since degradation by ClpXP is prevented.	M2	*C. crescentus* (a)	Joshi *et al* 2015

Table 1. Overview and summary of highlights and key information of published substrate trapping studies for the AAA+ proteolytic systems of Clp, FtsH and Lon families.

Bait	Key points and novelty	Technical aspects and other comments	affinity tags	Species and substrate pool	reference
Chaperone trap					
ClpC	About 100 proteins identified, including known Clp substrates and adaptors. Use inducible ClpCtrap in clpc null background . Subset also found in the *S. aureus* ClpPtrap study by Feng et al (2013). Most proteins are involved in stress response.	1D gel, MSMS, spectral counting, student t-test ($p<0.05$) (n=2). Additionally 12 proteins only in trap (p-value >0.05). Some with very low SPC (1-3). clpc/Empty vector as control; clpc/ClpCtrap for trapping. Inducible promoter. High number of proteins in trap (465, 634) and empty vector control (382, 521). Thresholds for candidate substrates include repeat observation and fold-enrichment. 8 proteins showed strong binding to affnity columns, likely due to their metal binding capability. Follow-up by e.g. bacterial two-hybrid analysis did confirm several interactions, but not neccesarily Clp degradation dependency.	His6-TEV-Myc	*S. aureus* (a)	Graham *et al* 2013
ClpC	*In vivo* chloroplast ClpC1 substrate trap with a C-terminal STREPII affinity tag identified a dozen specific interactors, as well as the ClpPRT protease core. Several of these trapped proteins over-accumulated in clp mutants and/or were found as interactions for the adaptor ClpS1, supporting their functional relationship to ClpP. ClpC1-TRAP induced a dominant visible phenotype.	*in vivo* ClpC1 substrate trap (in wt background) with a C-terminal STREPII affinity tag by mutating critical glutamate residues (E374A and E718A) in the two Walker B motifs of ClpC1 required for hydrolysis of ATP. Control is tagged ClpC1 in wt background. Thresholds for candidate substrates include repeat observation and fold-enrichment.	StrepII	*A. thaliana leaf* (d)	Montandon *et al* 2018
ClpX (*in vitro*)	Three putative mitochondria ClpX substrates, carbamyl phosphate synthetase (CPS), polymerase interacting protein of 38kDa (PDIP38) and p32 were identified by incubation of a ClpX WalkerB mutant protein with mitochondrial proteome. p32 could not be degraded by human ClpXP *in vitro* . It was speculated that p32 could be an adaptor for ClpX, or alternatively post-translational modification of p32 (e.g. by phosphorylation) is required to make p32 a substrate for ClpXP(P32).	*In vitro testing* of recombinant human ClpX Walker B domain mutant (glutamate to alanine) showed that this mutant version can still bind ATP, form ClpX oligomers and interact with the ClpP protease core. Mutant ClpX was incubated with mice liver mitochondrial proteome.	His6	human ClpX €	Lowth *et al* 2012

Table 1. Overview and summary of highlights and key information of published substrate trapping studies for the AAA+ proteolytic systems of Clp, FtsH and Lon families (continued)

Bait	Key points and novelty	Technical aspects and other comments	affinity tags	Species and substrate pool	reference
Protease trap					
ClpP	Breakthrough study establishing the concept of *in vivo* ClpPtrap with ClpX as chaperone. More than 50 proteins identified with range of functions. Confirmed in vivo trap using GFP-ssrA. Deletion analysis, fusion proteins, point mutations and peptide arrays identified several (tentative) N-terminal and C-terminal motifs recognized by ClpX.	Early days in proteomics (LC-Deca instrument)- using one replicate and with single peptide ids accepted. ClpAPtrap or ClpXPtrap for trapping and *clpA-clpX-clpP-* /ClpPtrap as negative control. ssrA-tagged substrates by deletion of smpB. Inducible promoter. Compared proteins identified in the presence of ClpX, ClpA, or both.	Myc3-TEV-His6	*E. coli* (a)	Flynn *et al* 2003
ClpP	SspB is an adaptor that delivers C-terminally ssrA-tagged proteins to the AAA+ protease ClpXP for degradation. ClpP trapping experiments showed that a fragment of RseA is delivered by SspB to ClpX for degradation, but surprisingly this was independent of any ssrA tag. Instead the RseA fragment has a C-terminus ending wih VAA, a C-motif 1 class of ClpX recognition signals. Thus SspB also delivers substrates to ClpX not tagged by ssrA. Follow-up experiments showed that SspB dependent degradation of RseA results in induction of the δE regulon.	*clpA-sspb+* /ClpXPtrap for SspB dependent substrates. *clpA-sspB-clpP-* /ClpXPtrap as control. 2D gel to compare the proteins enriched in the presence SspB. Focus on the main gel spot representing a fragment of RseA. Several follow-up experiments confirm RseA a ClpX-sspB substrate and the C-terminal VAA residues as the degron.	Myc3-TEV-His6	*E. coli* (a)	Flynn *et al* 2004
ClpP	Identified proteins degraded by the ClpXP system in response to DNA damaged by addition of chemical nalidixic acid (NAL) inhibiting, causing double strand DNA breaks and inducing the SOS response. Substrates include RecN and UvrA. Damage-response proteins appear an unusually rapidly degraded family and ClpXP has substantial capacity to process the influx of newly synthesized substrates while maintaining the ability to degrade its other substrates in an environmentally responsive manner.	SILAC used to identified conditionally enriched substrates of ClpXP after DNA damage. *clpA* -/ClpPtrap as trapping strain; *clpX-clpA-clpP-* /ClpPtrap and *clpX-* /ClpPtrap as controls. Inducible promoter. Experiment carried out both using constant time and similar cell density when comparing untreated and NAL treated cultures. Proteins enriched more than 3 fold after NAL treatments are highlighted. Several follow-up experiments.	Strep	*E. coli* (a)	Neher et al 2006
ClpP	About 70 proteins trapped in two different strains (Newman and 8325-4) and identified as candidate ClpCP, ClpXP or ClpXCP substrates, including transcriptional regulators CtsR and Spx, the ClpC adaptor proteins McsB and MecA, and the cell division protein FtsZ. Newly identified ClpP substrates include the global transcriptional regulators PerR and HrcA, proteins involved in DNA damage repair (RecA, UvrA, UvrB), and proteins essential for protein synthesis (RpoB and Tuf). This ClpPtrap study underscores the central role of Clp-proteolysis in a number of pathways that contribute to the success of *S. aureus* as a human pathogen.	Both clpp/empty vector and clpp/ClpPwt as control for clpp/ClpPtrap using an inducible promoter for the tagged Clp gene in 'wild-type' backgrounds or backgronds lacking either ClpX or ClpC. Inducible promoter. 2D gel used for separation of PTM varients and gel-free system to increase MS sensitivity for protein identification. The 2D gel did not provide evidence for differential PTMs in Clp substrates. Several follow-up experiments for Spx, TrfA, Cod4, Sle1 and RecA to determine protein stability and degradation, including inducton of the SOS response. It is pointed out that induction of the ClpPtrap results in protein stress, which can bias the outcome.	His6	*S. aureus* (a)	Feng *et al* 2013

Table 1. Overview and summary of highlights and key information of published substrate trapping studies for the AAA+ proteolytic systems of Clp, FtsH and Lon families (continued)

Bait	Key points and novelty	Technical aspects and other comments	affinity tags	Species and substrate pool	reference
ClpP	Identification of 32 high confidence substrate candidiates for ClpXP and specific protease recognition motifs. Degradation of stalk synthesis transcription factor TacA is controlled during the cell cycle dependent on the ClpXP regulator CpdR. Stabilization of TacA increases degradation of another ClpXP substrate, CtrA, while restoring deficiencies associated with prolific CpdR activity. This study shows new, validated ClpXP substrates, clarifies rules of protease substrate selection (*e.g* . C-terminal degron AA and SA, but also AG), and demonstrates how regulated protein degradation is critical for *C. crescentus* development and cell cycle progression. The study also points out that homologs of ClpXP substrates are also substrates in other species, in partcular *E. coli;* In other words there seems significant conservation of substrate-protease relationship across bacterial species.	Importantly, expression of the ClpPtrap in presence of wt Clp levels fails to trap substrates, suggesting that partial reduction of the number of catalytic triads per Clp core does not significantly result in stalled degradation. However, succesfull trapping when ClpPTRAP is constitutively while native ClpP is down-regulated resulted in overaccumulation of substrates in the trap. Tagged Clp-WT was used as negative control. Whole cell extracts used as the reference of the cell abunandant proteins. Application of high stringency thresholds for substrates (minimal 5 unique peptides in both replicates and not copurified from control and not abundant proteins in whole cell extracts).	His6-TEV-M2Flag	*C. crescentus* (a)	Bhat *et al* 2016
ClpP	Identification of ~20 strong candidate substrates of ClpXP traps, mainly associated with metabolic pathways in mitochondria, e.g. pyruvate dehydrogenase complex, TCA cycle, complex I, and amino acid and fatty acid metabolism. This suggests that the main function of ClpXP in the mitochondrial matrix is in the control and/or maintenance of mitochondrial (energy) metabolism. Note that deletion of ClpP in this fungal species results in increased healthy life-span.	The ClpPtrap protein is blocked in autocatalytic N-terminal processing and cannot form a normal Clp core assembly. Therefore, the study used a human mitochondrial ClpPtrap to complement the null mutant, using the human active ClpP and the ClpP null mutant as controls. Funtionally complementing using HsClpPtrap as alternative. Based on a previous study, endogenous ClpX appears to normally interact with the human ClpP. It is noted that human mitochondrial ClpP does not undergo N-terminal autocatalytic processing.	Flag3-His6	*P.anserina* (c)	Fischer *et al* 2015
ClpP	Mammalian ClpP indirectly regulates mitoribosome assembly and thereby protein translation. ClpXP is needed to remove mitochondrial ERAL1 (an Era GTPase), a putative 12S rRNA chaperone and mitochondrial ribosomal assembly factor. Timely removal of ERAL1 from the small ribosomal subunit is essential for the efficient maturation of the mitoribosome and a normal rate of mitochondrial translation. Other candidate mammalian ClpXP substrates are involved in the regulation of mitochondrial translation, oxidative phosphorylation, and metabolic pathways. Overlap between mammalian ClpXP and *E. coli* ClpXP substrate homologs, *i.e.* significant conservation of substrate-protease relationship across bacteria and mitochondria.	Compare clp null mutant (*clpp-*) with tagged ClpPwt and ClpPtrap overexpressed in *clpp-* cells for 48 hr under constutive promoter before harvest. Observed stabilized interaction between ClpP and ClpX, likely due to bloked degradation. Follow-up studies indicate some candidates are likely ClpX interactors (P32/C11QBP), others (EFG1 and ERAL1) are indeed substrates, but yet several others are neither. Clearly important to follow-up with degradation analysis	Flag	Mouse (f)	Szczepanowska *et al 2016*
ClpP	Discovery of arginine phosphorylation as degradation signal for ClpCP in *B. subtilus* using *a* ClpP-trap and MSMS and many controls and follow-up experiments. *In vitro* Arg phosphorylation by McsB kinase is required and sufficient for the degradation, and counteracted by arg phosphatase YwlE. The docking site for P-Arg is located in the amino-terminal domain of the ClpC ATPase, as resolved at high resolution in a co-crystal structure. This appears specific for gram positive bacteria and absent in gram negative bacteria (*e.g. E. coli*). Recruitment of arg-P substrates by ClpC does overcome the need for MecA in ClpCP assembly. Parallels pointed out to polyubiquitination.	Heat stress activated the ClpC adaptor actvity and potential substrate pools before trapping. Avoid mixing of ClpPtrap and endogenous ClpP *in vivo* by point mutations (see also Bhat et al 2016). wt/ClpPX as control for wt/ClpPXtrap. Additonal *clpC-*/ClpPX as control for *clpC-*/ClpPXtrap. Thresholds for candidate substrates include repeat observations and fold-enrichment. Extensive follow-up experiments, incluiding structural analysis of ClpC-N-domain.	His6	*B. subtilis* (a)	Trentini *et al* 2016

Table 1. Overview and summary of highlights and key information of published substrate trapping studies for the AAA+ proteolytic systems of Clp, FtsH and Lon families (continued)

Bait	Key points and novelty	Technical aspects and other comments	affinity tags	Species and substrate pool	reference
ClpP3 and ClpP5	The chloroplast Clp protease core complex contains 4 different catalytically inactive ClpR and 5 different CLPP homologs (P1,3-6), and two ClpT proteins, and total of 10 catalytic triads. ClpP3trap complements clpp3 null mutant, but ClpP5trap did not complement clpp5 null. Comparative of MS/MS showed that ClpP3trap, ClpP5trap, ClpP3WT or ClpP5WT did not suggest trapped substrates, suggesting that the bottle-neck for degradation is substrate recognition/unfolding by ClpP adaptors and chaperones, upstream of the ClpP core. This agreed with ClpPtrap study in C. cresentus (Bhat et al 2016) (and see Trentini et al 2016) showing that heterooligomers of ClpP and ClpPtrap core do not accumulate substrates.	Because loss of Clp proteolys is embryo lethal, a fully inactive ClpPtrap can not be generated. Endogenous promoters and heterozygous lines were used to overcome lethality in the ClpP3/5 lines. Stringent thresholds for candidate substrates include mulitiple unique peptides, repeat observations and minimum fold-enrichment. Endogenous biotin-conjugating proteins identified as contaminating proteins on the streptavidin column.	Strep	*A. thaliana* (d)	Liao *et al* 2018
FtsH	The first published FtsHtrap is based on inactivation of proteolytic site, but with (confirmed) intact unfolding and translocation activity of substrates into the inactivated proteolytic chamber. Out of 14 candidate trapped substrates, four novel FtsH substrates were confirmed in follow-up experiments, namely IscS, DadA, FdoH and YfgM which serves as negative regulator of the RcsB-dependent stress response. FtsH-dependent degradation of YfgM pathway facilitates this stress response. This extends the number of *E. coli* FtsH confirmed substrates to 15.	Control experiments with purified FtsH variants confirm activities. FtsH proteolytic trap accumulated the known substrate LpxC. *ftsh*-/FtsHwt as negative control for *ftsh*-/FtsHtrap. Inducible promoter. 2D gel was used to compare control and trap. Not all trapped are substrates.	His6-MBP	*E. coli* (a)	Westphal *et al* 2012
FtsH	Follow-up & extension of Westphal *et a* l 2012. Label-free nanoLC-MSMS of FtsHtrap (inactive proteolytic site, normal unfolding) and FtsHwt interactors *in vivo* from exponential and stationary growth stages. Of >50 proteins identified including known substrates; 4 (SecD, ExbD, YlaC, YhbT) out of 37 tested new candidates were confirmed substrates through *in vivo* degradation. Six others were degraded by one or more other protease(s). Known YfgM degron in N-terminus appears also in ExbD and YlaC suggesting a general degron (acidic-polar-hydrophobic-acidic) for FtsH.	Similar genetic material as Westphal et al (2012) but now cells from different physiological stages and used label-free LC-MSMS for identification and quantification (not 2D gels). Quantification by peak intensity. Clearly stated criteria for substrate assignment. Follow-up *in vivo* degradation experiments. Short discussion and references to degrons for AAA+ proteases. For detailed protocol see Lindemann *et al* 2018.	His6-MBP	*E. coli* (a)	Arends *et al* 2016
FtsH4	The first substrate trapping effort for a plant mitochondrial protease FtsH4 in the inner membrane. Comparison between control purification with sample from the *fths4* null and FtsH4-trap-flag expressed in the *ftsh4* null. Some 17 proteins were 2-fold enriched based on spectral counts and include Tim17-2, the only known FtsH4 substrate. Other candidate substrates include MPC4 and Pam18-2; follow-up in organello degradation essays provides support for these candidate substrates.	ftsh4-1/FtsH4trap for trapping and *ftsh4-1* without any tagged transgenic protein as control. MS analysis provides a copurified protein list. The lack of Ftsh4-WT-flag control to determine enrichment in the Ftsh4-trap-flag makes it harder to interpret the significance of the identified proteins.	Flag	*A. thaliana* (c)	Opalińska *et al* 2017
Lon	Parallel comparative (SILAC) proteomics (lon+ and lon-) and Lontrap (proteolytic inactive) approach combined with MS/MS discovered Lon-dependent substrates and physiological functions. Many enriched candidate substrates identified, but ultimately 14 were *in vivo* confirmed novel Lon substrates involved in *e.g.* superoxide stress response, sulfur assimilation, nucleotide biosynthesis, amino acid and central energy metabolism. Interesting discussion on why not all trapped proteins could be confirmed as in vivo substrates with the assays used, and how comparative proteomics and trapping approaches are complementary.	Control experiments with purified Lon variants confirm ATP-hydrolysis and inactivation proteolysis. *lon*-/Lonwt as control for lon-/Lontrap for trapping; SILAC analysis of lon-/empty vector or lon-/Lonwt as well as Lon+ and Lon++ (overexpression). Inducible promoter. SILAC (global view of the proteome with/out lon) provides insight in lon contributions to cellular homeostasis. Quantification by peak intensity and clearly stated criteria for substrate assignment; Vulcano plots. Follow-up *in vivo* degradation experiments.	His6	*E. coli* (a)	Arends *et al* 2018

The proteome used was soluble cell extract (a), chloroplast stroma (b), mitochondria (c), soluble leaf extract (d), mitochondria in mice liver (e), or embryonic fibroblasts (f)

proteins and peptides. The main challenge and drawback of these comparative proteomics approaches is that it is often difficult to distinguish between primary and secondary (pleiotropic) effects caused by induction or loss of protease activity. For instance, in the case of the well-studied plant and algal chloroplast FtsH and Clp loss-of function mutants, strong phenotypes result in a wide range of molecular changes, many of which are pleiotropic (reviewed in (*24-28*)). Even using transient (inducible) loss of protease capacity can result in pleiotropic phenotypes due to the relative stability of the targeted protease (*29*). Nevertheless, these proteomics approaches help to define protease functions, and can result in candidate substrates, but more targeted follow-up studies are needed to further test if indeed these candidates are *bona fide* substrates (see BOX 3).

An effective mass spectrometry (MS) -based approach has been developed to determine cleavage specificity of proteases using recombinant protease and peptide libraries generated from (sub)cellular proteomes. This technique is called PICS (proteome identification of cleavage sites) and involves differential stable isotope labeling of the protease-treated peptide libraries and non-treated peptide libraries, allowing identification of peptides unique to the protease-treated sample. These unique peptides then likely represent the cleavage products of the protease, and when combined with predicted proteome information, protease cleavage site specificity can be visualized in so-called sequence logos (https://weblogo.berkeley.edu/logo.cgi) and potential substrates can be identified (*21, 30, 31*). The PICS strategy has been shown to be highly effective for a wide range of proteases (e.g. serine, metallo, soluble, membrane-bound) that display a narrow or defined cleavage specificity, *e.g.* (*32-34*).

However, proteases that are processive, such as the members of the AAA+ protease families, generally lack cleavage specificity (but see (*35*)), and generally have no or little preference for cleaving peptide bonds of specific amino acids, rendering the PICS technique less informative.

2.4 *IN VIVO* SUBSTRATE TRAPPING

In vivo trapping can overcome the limitations of the proteomic approaches discussed above, and is an increasingly popular approach for substrate identification. Table 1 lists and summarizes key information of all (to our best of knowledge) publications using substrate trapping techniques for members of the ATP-dependent AAA+ protease families. The concept of *in vivo* trapping is to capture substrates through their interaction with components of a protease system, essentially by preventing degradation, and stabilizing the interaction between substrate and protease system component (Figure 1). This is then followed by affinity enrichment of the tagged protease component and its interactors, typically followed by MS-based identification (Figure 2). Several prerequisites are required to identify substrates using *in vivo* trapping: (i) proteolysis is prevented or strongly delayed for the targeted substrates, (ii) the physical interaction of the trapped substrate(s) and the proteolytic machinery must be stable enough to allow co-purification, and (iii) a parallel negative control is required to recognize unspecific interactions with the protease component, to the affinity matrix, and other contaminating proteins (*e.g.* high abundant proteins unrelated to the protease).

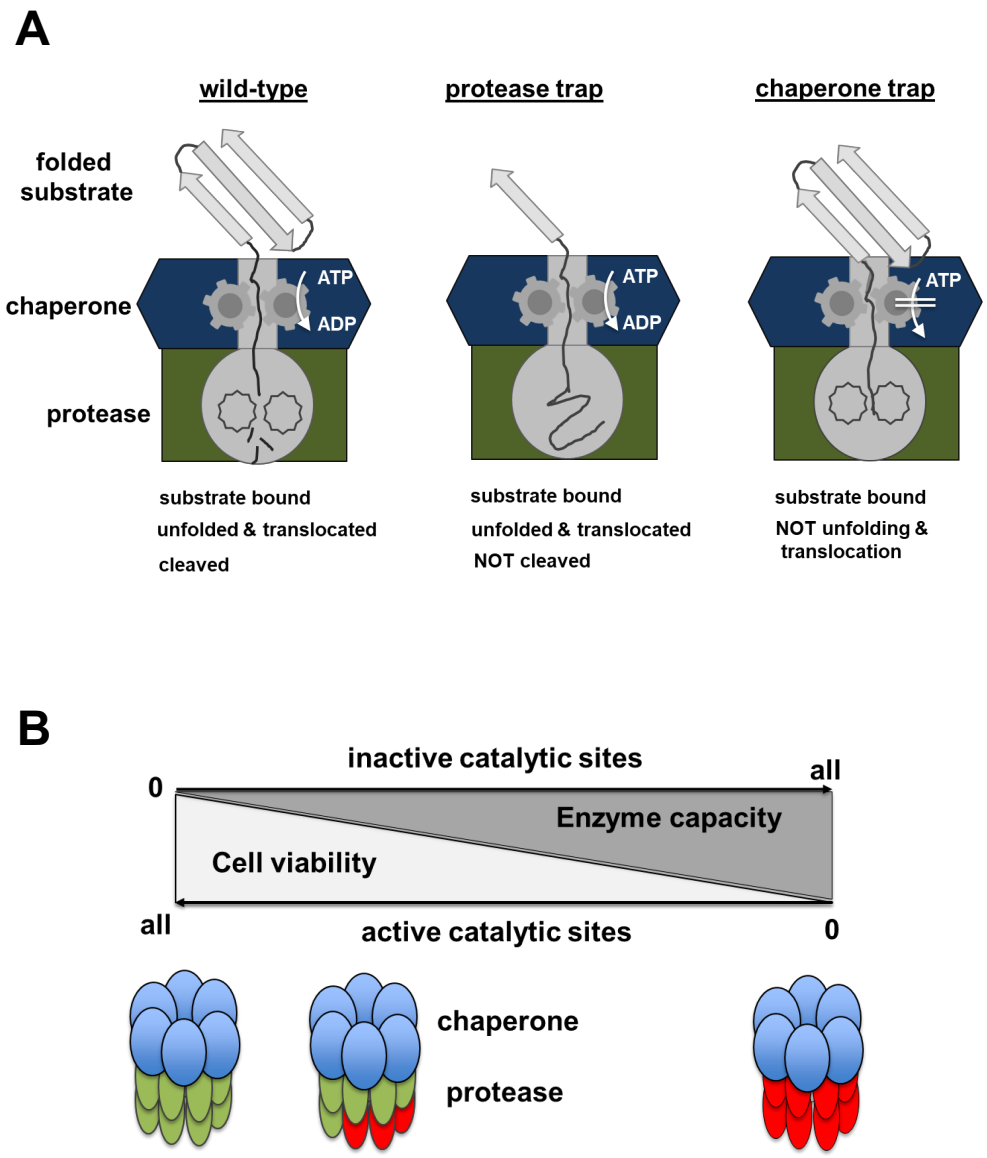


Figure 1. Schematic outline of in vivo substrate traps. A, Cartoon of control trap, protease substrate trap (defective in proteolytic triad) and chaperone trap (defective in ATP hydrolysis and unfolding). **B,** In case of proteases that affect cell viability, an increase in the copy number of inactive protease results in reduced viability.

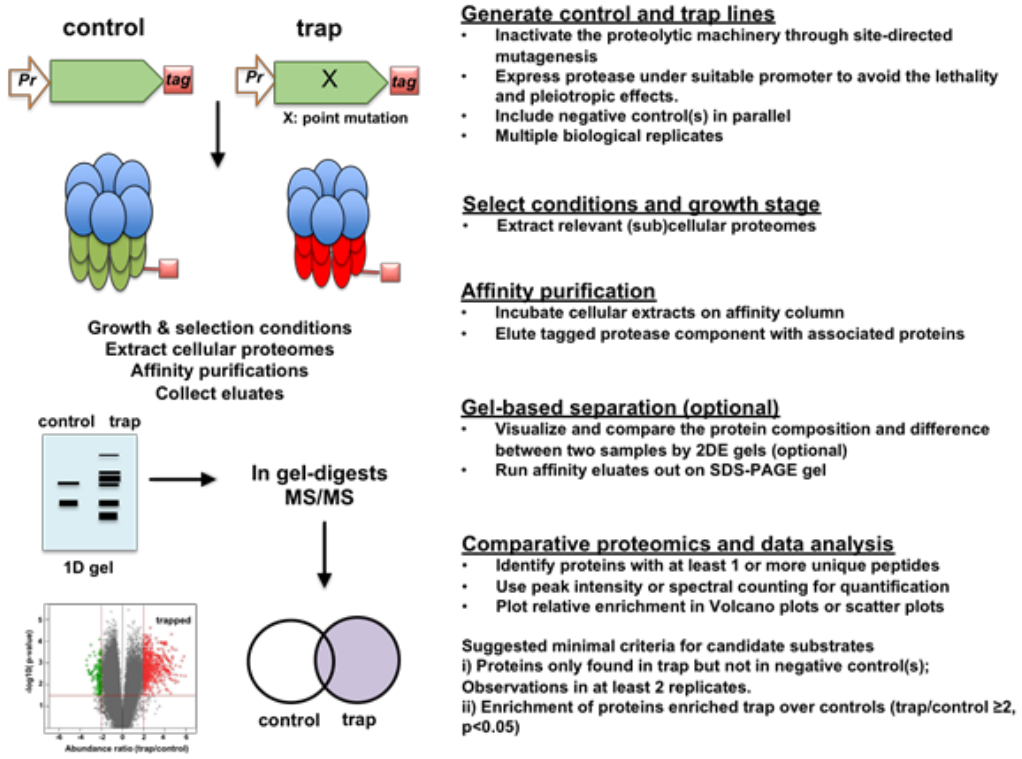


Figure 2. The workflow of *in vivo* substrate trapping.

2.4.1 *In vivo* trapping using catalytically inactive proteases

In this approach, substrates are trapped within proteolytic chambers by blocking peptide bond cleavage through site-directed mutagenesis of catalytic sites (Figure 1). This approach was pioneered by the Baker lab for the *E. coli* ClpP protease, in which the catalytic serine was replaced by alanine (*36-38*). This point mutation (S98A) inactivates proteolytic cleavage within the ClpP complex and traps the substrates. Using this trapping approach in different genetic backgrounds, these studies probed for the role of chaperones ClpX and ClpA in substrate selection, the role of SspB in C-terminal SsrA tagging of substrates, and also the physiological role of Clp in removal of proteins following DNA damage. Importantly, this also helped to

define degrons recognized by the Clp system. This Clp protease catalytic trapping technique has since been applied to Clp proteases in *S. areus, C. crescentus, B. subtillus*, *P. anserine* mitochondria, mouse mitochondria and *A. thaliana* chloroplasts (Table 1). These studies provided insight in the role of Clp in this range of species and lead to the exciting discovery of arginine phosphorylation as degron for ClpCP in *B. subtilus* (*39*). However, these studies also demonstrated several challenges using these catalytically inactive ClpP proteins, including problems in assembly of the Clp core complex due to interference with auto-catalytic N-terminal processing of ClpP, and loss-of viability phenotypes (Table 1 and see below section '***Challenges in generating substrate-trapping lines'***).

There have been several *in vivo* studies with inactivated proteolytic traps for FtsH in *E. coli* (*40, 41*), FtsH4 for *A. thaliana* mitochondria (*42*), and for Lon in *E. coli* (*43*). In both FtsH and Lon, the protease and unfoldase domains are within the same protein, which is different to Clp system in which the protease and unfoldase are encoded by different genes. Importantly, it was shown that the catalytic peptide bond cleavage activity can be abolished through a point mutation in the Zn^{2+}-binding site of FtsH, or in the catalytic serine of Lon, while the ATPase activity of the unfoldase domain remained intact to ensure that substrates are translocated into the proteolytic domain and become trapped. These FtsH trapping studies identified candidate substrates, several of which were confirmed by follow-up degradation studies (*40*) (see also BOX 3). The application of inducible promoters was important to avoid pleiotropic effects in the bacterial studies (Table 1) as well discussed further below.

2.4.2 *In vivo* trapping using inactive AAA+ chaperones or AAA+ chaperone domains

A complementary approach to catalytic protease traps is to trap by preventing the release of substrates from the AAA+ chaperone (domain) (Figure 1). For Clp proteases, substrate unfolding and translocation depend on separate Clp chaperones A, X, C, or D (*11, 13*), which are members of the larger family of HSP100 proteins that include the well-studied ClpB and HSP104 disaggregases/foldases in bacteria, yeast, and plants (*44-46*). Each of these chaperones have either one or two AAA+ domains, which contain the Walker A and B domains involved in ATP binding and hydrolysis (*13*). The interaction of substrates with AAA+ chaperones is ATP-dependent, whereas the subsequent unfolding and release of substrates requires ATP hydrolysis. Substrates of the *E. coli* ClpB (not coupled to ClpB proteolysis) remain stably associated when ATP hydrolysis was blocked through mutagenesis of a critical glutamate residue in the Walker B domains, allowing for identification of *in vivo* substrates (*47, 48*). This *in vivo* substrate trapping concept was applied to ClpC in *S. aureus* (*49*) and ClpC1 in *A. thaliana* (*50*) (Table 1). The *S. aureus* study used an inducible ClpC-TRAP in a *clpc* null background, resulting in identification of some 100 candidate substrates, including known Clp substrates and adaptors, several of which were also found in the ClpP protease trap (*51*). To our knowledge chaperone domain traps have not yet been reported for FtsH or Lon family members.

2.5 CHALLENGES IN GENERATING SUBSTRATE TRAP LINES

Completely blocking AAA+ proteases *in vivo* can result in lethality or other pleiotropic phenotypes (*25, 39, 52*). To overcome these challenges and obtain physiologically relevant results, trap lines were created in which either the transgenic trap protein was induced or the wild-type protein was down-regulated (inducible depletion) (Figure 1B) This inducible strategy has successfully been used in *E. coli* (ClpP, FtsH, Lon), *S. aureus* (ClpC, ClpP) and *C. crescentus* (ClpP*)* (Table 1). An IPTG-inducible promoter to express inactive ClpP *in vivo* was used to avoid the pleiotropic phenotype of ClpP deletion in *B. subtilis* (*39*). This was combined with mutagenesis of specific residues (E119R/R142E) in ClpP to generate electrostatic repulsion between the recombinant and endogenous ClpPs *in vivo*, preventing formation of a heteromeric complex of endogenous ClpP and recombinant ClpP-trap proteins. This strategy allowed expression and enrichment of tagged-ClpP complexes at selected conditions. An alternative approach was used for the essential ClpP in *C. crescentus* in which the wild-type ClpP under control of a xylX promoter was depleted, while tagged active (as negative control) or inactive ClpP trap proteins were constitutively expressed (*52*). Both these approaches avoided mixing of active and inactive ClpP in the same complex; such mixed complexes might be ineffective substrate traps, especially if the rate-limiting step for substrate selection, unfolding and degradation is upstream of the proteolytic events. Indeed, in case of the highly heteromeric chloroplast ClpPR complex in *A. thaliana*, tagged complexes with inactive ClpP3 or inactive ClpP5 did not result in obvious trapped substrates (*53*).

2.5.1 The choice of affinity tags

Different affinity tags have been employed in substrate trapping studies. These tags are either placed at the N- or C-terminal end. In case of proteins that have a (cleavable) N-terminal signal sequence (*e.g*. for sorting to mitochondria or chloroplasts) the tag is placed at the C-terminus as not to interfere with the sorting process. ClpP in *E. coli* (and likely many other species) (*54*), the Lon protease PIM1 in mitochondria of *S. cerevisiae* (*55*) and mitochondrial FtsH members (m-AAA+ proteases) (*56*) undergo N-terminal autocatalytic processing. Such processing is another reason to avoid placing the affinity tag at the N-terminus. The affinity tags used in the studies listed in Table 1 include the single tags polyhistidine (His6) and StrepII, as well as tandem tags with cleavage sites (Myc3-TEV-His6, His6-TEV-M2FLAG) or without cleavage sites (Flag3-His6, His6-MPB). The choice of the affinity tag and associated affinity matrix is influenced by: i) yield, ii) unspecific interactions to the affinity matrix, iii) ease of use and costs, for *e.g*. the affinity matrix, and iv) availability of antiserum against the affinity tag. Based on repeated affinity purifications and systematic MS analysis (see BOX 1), lists of typical contaminants can be generated – these lists are very useful for the experimenter to recognize potential false-positive interactors. Recognition of contaminants and non-specific interactions is a key reason why suitable negative controls are needed in trapping experiments (see BOX 2). There is a database with contaminant data from affinity purifications (not specific to proteases) entitled CRAPome (Contaminant Repository for Affinity Purification Mass Spectrometry Data) at http://www.crapome.org/ current only for *E. coli, H. sapiens* and *S. cerevisiae* (*57*). Community-wide contributions to this or similar databases would be helpful for future trapping studies.

2.5.2 Conditional enrichment of substrates

Because a key role of proteases and proteolysis is to respond to changes in developmental state, the (a)biotic environment or other temporal changes, substrates are sometimes only present (or in higher abundance) under specific conditions. When systematically searching for protease substrates, it is therefore important to consider such temporal conditions. It is worth noting that some conditions may have a stronger effect on the trap lines than the control lines, which seems intuitive as proteases support fitness. As an example, using SILAC-based (Stable Isotope Labeling by Amino acids in Cell culture; see Glossary) quantitative comparative proteomics, it was found for *E. coli* Clp trap lines that differential cell growth and cell density must be considered when comparing control and trap lines (*38*) (see also BOX 2). Similarly, proteases or components of proteolytic systems, such as chaperones or adaptors, might be expressed in particular cellular states or only upon specific environmental conditions. It is therefore also important to consider the choice of promoter that drives the transgenic affinity-tagged trap. Endogenous (genomic) promoters would perhaps yield the most physiologically relevant results, but low expression rates or practical considerations (*e.g.* cloning issues) could be an impediment to choose such endogenous promoters. Comparing protease trap lines in backgrounds that lack other components of a protease system, such as adaptors (*e.g.* ClpS, SspB) or specific chaperones (ClpX, ClpA, ClpC) allows determination of substrate selection mechanisms. This has successfully been used in several studies listed in Table 1, *e.g.* (*36, 37, 51*).

2.5.3 *In vitro* trapping using cellular extracts

In vitro trapping of cellular or subcellular proteomes extracted from different genetic backgrounds or after treatments of organisms by specific stresses is an excellent complement for systematic substrate identification of proteolytic systems. It is complementary because i) there is no need to generate an *in vivo* trap which can be challenging (or even impossible) and/or time-consuming, and ii) it allows to rapidly test the significance of specific domains or amino acid residues in the protease component. However, the main limitation is that substrates must be able to accumulate *in vivo* prior to extraction of the (sub)cellular proteome, whereas in the case of *in vivo* traps, substrates can accumulate in the trap over time. There are three examples this *in vitro* strategy; in all cases this concerned the adaptor ClpS in either *E. coli* (*6, 58*) or in chloroplasts from *A. thaliana* (*59*) (Table 1). In all three studies the specificity of protein interactors to ClpS was determined by using a comparison of ClpS and ClpS mutated in two residues involved in N-degron recognition (*E. coli* ClpS-D35A/D36A and *A. thaliana* ClpS1-D89A/N90A). The chloroplast study and the initial *E. coli* study (*6*) extracted cellular proteomes of a mutant deficient in Clp chaperone capacity. In case of the *E coli* study, the mutant was devoid of ClpA but the ClpX homolog was present, allowing identification of substrate specifically depending on ClpA substrates (since only these substrates overaccumulated). In case of *A. thaliana*, the plants were null mutants for chloroplast ClpC1, a functional homolog of ClpA, whereas the other chloroplast Clp chaperones ClpC2 and ClpD were both present. ClpC1 is the dominant chaperone, since null alleles for ClpC2 and ClpD do not show visible phenotypes, whereas ClpC1 null alleles have reduced growth and virescent phenotypes (*59-61*).

The initial, ground-braking *E. coli* study resulted in the important discovery that one of the ClpS substrates, putrescine aminotransferase (PATase), is post-translationally modified by a leucyl/phenylalanine tRNA-protein transferase (LFTR also named AAT) to generate an N-degron (*6*). This PTM is not only essential for its recognition by ClpS, but also determines the stability of PATase *in vivo*. It is important to note that prior to becoming a ClpS substrate, PATase was N-terminally cleaved by unknown protease activity. The second *E. coli* study (*58*), using cell extracts from stationary phase cells and elution with a peptide bearing an N-degron, identified more than 100 specific ClpS interacting proteins. Most of these ClpS interactors were N-terminally truncated by unidentified protease activity, and many were further modified presumably by AAT. Importantly, Edman degradation sequencing showed that 32 out of 37 N-termini had N-degrons, in particular Leu and Phe. This *E. coli* study demonstrated that widespread proteolytic processing of cellular proteins by unknown protease activity did generate N-end rule substrates through the action of an amino-transferase and subsequent selection by the ClpS. In case of the chloroplast study, eight high confidence candidate ClpS1 substrates were identified, including glutamate-tRNA-reductase (GluTR) which is a key control point in tetrapyrrole synthesis (*59*). Most of these candidate substrates were also found over-accumulating in loss-of function mutants for ClpC1 and/or ClpS1, as determined by comparative proteomics using spectral counting (see Glossary – label free quantification). Finally, there was one protein, ClpF, which strongly interacted with both ClpS1 and ClpS1-D89A/N90A, suggesting it is not a substrate. Follow-up studies

suggested that ClpF is unique co-adaptor of ClpS1 (*62*) and showed that GluTR degradation is delayed in Clp mutants (*63*).

2.6 CONTROLS USED FOR TRAPPING STUDIES

To be able to recognize strong candidate substrates among the proteins identified in affinity eluates of *in vivo* or *in vitro* traps, it is essential to have suitable negative controls. These negative controls should closely mimic the trap except for the residues that were modified to trap the substrates. Indeed, most studies listed in Table 1 use unmodified affinity tagged controls. However two of the studies lack negative affinity-tagged controls and used either an empty vector control or just the null mutant for the protease (*42, 49*). These empty vector controls allow recognition and removal of contaminants interacting with the affinity matrix, or proteins simply present due to their very high abundance in the original cellular extract (see also the section above '**The choice of affinity tags'**). However these empty vector controls cannot recognize proteins interacting with the tagged protein for reasons unrelated to being a trapped substrate. This is a concern, because AAA+ chaperones (ClpA, C, X) or chaperone domains in Lon or FtsH proteins have a natural tendency to interact with proteins, especially when these proteins expose a more hydrophobic region, *e.g*. due to partial unfolding or protein damage.

Negative controls for key domains or residues that are important for substrate selection or interaction with upstream components can be very effective to further establish if proteins are likely substrates (*e.g.* the residues in ClpS and ClpS1 that provide specificity to N-degron recognition and binding (*6, 58, 59*)). An extensive list

of considerations and criteria for confident candidate substrate identification is provided in BOX 2.

2.7 CONCLUDING REMARKS AND FUTURE PERSPECTIVES

In vivo and *in vitro* trapping, affinity purification and MS analysis of putative protein interactors co-purified with proteolytic systems, provides a systematic and unbiased way to study protease substrates, interactors, adaptors, and degrons. This approach has successfully been applied to both the soluble (Clp and Lon) and membrane-bound (FtsH) proteases in the AAA+ protease families. This has resulted in elucidation of protease contributions to proteostasis, as well as substrate selection pathways, degrons and protein modifiers. However, much remains to be learned about (conditional) substrates and selection mechanism for the AAA+ protease families, including the specific roles of protease orthologs, in particular in the multigene families in organelles of eukaryotes. We hope that the summary and evaluation of published substrate trapping studies in this review will inspire a broader use of this approach, including for proteolytic systems outside of the AAA+ families.

2.8 HIGHLIGHTS

1. *In vivo* trapping has been highly effective to discover substrates for AAA+ proteases Clp, FtsH and Lon, with most wide-spread use and success for the Clp system.

2. *In vitro* trapping with the adaptor ClpS using cellular proteomes from different genetic backgrounds discovered many (candidate) substrates, N-degrons for ClpS

substrates and an amino-transferase (LFTR or AAT) involved in generating N-degrons

3. *In vivo* and *in vitro* substrate trapping in AAA+ protease systems has enormous potential to better understand their contribution to proteostasis across many species, environmental conditions, different developmental states and genetic backgrounds.

2.9 OUTSTANDING QUESTIONS

1. What are the degrons of the trapped substrates for AAA+ proteases across different species? How do these degrons trigger substrate recognition and subsequent proteolysis? What internal or external signals trigger degron formation? Is there any hierarchy among degrons?
2. AAA+ proteases systems evolved and diversified. To what extent is this diversification an adaptive response to evolving (sub)cellular proteomes and environmental conditions?
3. AAA+ adaptors have not been used as *in vivo* traps, even if these could provide novel insights in the functional role and physiological substrates of adaptors. Can adaptors be modified to generate *in vivo* substrate traps?
4. Can the *in vivo* trapping approach be used for non-AAA+ proteolytic systems?

APPENDIX

BOX 1. The importance of mass spectrometry (MS) for substrate identification in protease traps Since the late 1990, MS has become the method of choice for protein identification, modification and quantification (*64-66*). Indeed, all studies listed in Table 1 used MS, in one case complemented by Edman degradation sequencing for determination of N-termini (*6*). Substrate trapping approaches combined with MS and various quantification techniques (*e.g*. SILAC, spectral counting), now allow for identification of ever larger numbers of candidate substrates. Consequently, this requires more rigorous statistical significance analysis and the need for larger numbers of biological replicates and appropriate controls. Furthermore, the advances in MS is shifting the bottleneck to follow-up studies in which the biological significance of the candidate substrates are tested. Table 1 includes comments on significance thresholds and follow-up studies.

BOX 2. Criteria for candidate protease substrate identification by trapping approaches Candidate substrates are expected to only be present or to be highly enriched in the trap but not the negative controls. We recommend five complementary considerations for confidence in identification of substrates:

1. *MS/MS identification criteria*. For confident identification within replicates, a minimum of two non-redundant peptide sequences observed with highly significant MS/MS ion scores is recommended.
2. *Identifications in independent replicates*. Candidate substrates should be observed in at least two independent replicates, but for more robust statistical analysis three or more replicates is strongly recommended.

3. *Suitable negative controls.* As a minimum, a negative control is essential to recognize interactors to the affinity matrix and unspecific interactors to the tagged protease component. Independent replicates are needed for robust statistical analysis, as in #2.
4. *Other controls.* Additional controls are strongly recommended, *e.g*. using different genetic backgrounds (*e.g*. loss of function mutants in Clp chaperones or the ClpS adaptor for ClpP protease traps) or different environmental conditions and/or developmental transitions or states. These additional controls will allow testing i) substrate delivery pathways (*e.g*. ClpX vs ClpA), ii) dependencies on adaptors, or iii) generation of degrons that turns proteins into protease substrates (*e.g*. SsrA, N-degrons).
5. *Quantification and statistical significance.* Relative protein enrichment of candidate substrates in the trap compared to the negative control can be determined by quantification based on the number of spectral counts and/or the peak intensity (see Glossary – label free quantification), possibly aided by differential stable isotope labeling. This enrichment can be visualized as a Volcano plot, plotting fold-change (ratio) against the statistical significance (p-value). For proteins identified only in either the sample or the control, missing values need to be replaced (imputed) with a value, typically the minimal value observed across the experiment. Statistical significance of enrichment can be tested using ANOVA or pairwise student t-test (with multiple hypothesis correction).

BOX 3. Follow-up and validation of potential substrates It is important to carry out follow-up experiments to determine the significance of proteins identified in traps. Several follow-up approaches are listed below:

1. Verify if homologs of the candidate substrates are known to be substrates in the homologous protease system in other species. Conservation of substrate-protease relationships would provide indirect support for a trapped protein being a substrate. Several examples of such conservation is pointed out in Table 1.
2. Determine if candidate substrate are over- or under-accumulating in mutants in the proteolytic system, as in (*43*) and (*50*). Over-accumulation of proteins provides indirect support for being a substrate.
3. *In vivo* degradation essays of the candidate substrate to test if the candidate is indeed degraded in dependence of the proteolytic system. This test can be done for endogenous proteins and tracked by immunoblots or targeted proteomics, or done using a specific (inducible) transgenic candidate substrate-reporter construct that includes a specific tag for detection. These tests can be carried out in wild-type or different genetic backgrounds, or under specific conditions, in particular mimicking those used for the initial trapping study. In some cases, inhibition of *de novo* protein expression (*e.g*. by addition of specific antibiotics) may be needed to determine the net degradation rate.
4. *In vitro* degradation essays using purified endogenous or reconstituted proteolytic systems added to the candidate substrates. Degradation is tracked using immunoblotting or other types of essays. Competition essays using known substrates as the competitor can be employed in this context.

5. Protein-protein interaction essays between candidate substrate and protease component such as yeast two hybrid system, bimolecular fluorescence complementation (BiFC), Fluorescence resonance energy transfer (FRET), etc. However, these approaches can easily result in false positive and false negative results. False positive results can occur because the protease component and substrate are normally not present at the same time or in the same subcellular location. False negative results can be due to the lack of a degron that is only induced under specific conditions or by specific enzymes.
6. Identification of the degradation signals shared by the trapped substrates. The recognition of substrates for each proteolytic system is key to dissect the protease web *in vivo.*
7. Test the putative degrons by deletion or mutations or by fusing degrons to otherwise stable reporter proteins.
8. Reciprocal trapping of the protease using the substrate with and without the putative degron to test the interaction between degrons and the proteolytic system.

GLOSSARY

1. *AAA+ protease system*: A specific group of ATP-dependent proteases that need ATP to unfold and translocate substrates into barrel-like proteolytic chambers for peptide bond cleavage. Clp, FtsH and Lon proteolytic systems are most studied examples.
2. *Degron:* A degron is a portion of a protease substrate that marks the protein for selection and degradation. Degrons include N- and C-terminal amino acid residues or motifs, as well as internal regions that act as degrons when exposed to the surface.
3. *Adaptor::* Proteins that recognize specific degrons and deliver these degron-marked proteins to a proteolytic system. Some adaptors also function in chaperone assembly.
4. *Comparative quantitative proteomics and peptidomics*: Techniques to systematically compare the protein composition and abundance among samples using mass spectrometry.
5. *N-terminomics/degradomics*: Systematic identification of N-termini using (isotopic) N-terminal labeling and enrichment strategies that allow determination of N-terminal processing and degradation events. The most popular techniques are TAILS and COFRADIC and variants thereof.
6. *Label free quantification:* A MS-based protein quantification method using peptide spectral counting or spectral peak intensity to calculate protein abundance in complex proteomes.

7. *SILAC:* Stable isotope labeling with amino acids in cell culture that allow comparative protein quantification. Usually two type of isotopes (heavy and light) are used to differentially label proteomes (*e.g.* with and without a specific protease), followed by mixing and MS-based comparative protein abundance determination.

REFERENCES

1. J. Labbadia, R. I. Morimoto, The biology of proteostasis in aging and disease. *Annual Review of Biochemistry* **84**, 435-464 (2015).
2. N. D. Rawlings, A. J. Barrett, R. Finn, Twenty years of the MEROPS database of proteolytic enzymes, their substrates and inhibitors. *Nucleic Acids Research* **44**, D343-350 (2016).
3. M. Rape, Ubiquitylation at the crossroads of development and disease. *Nature Reviews: Molecular Cell Biology* **19**, 59-70 (2018).
4. A. Santner, M. Estelle, The ubiquitin-proteasome system regulates plant hormone signaling. *Plant Journal* **61**, 1029-1040 (2010).
5. R. D. Vierstra, The ubiquitin-26S proteasome system at the nexus of plant biology. *Nature Reviews: Molecular Cell Biology* **10**, 385-397 (2009).
6. R. L. Ninnis, S. K. Spall, G. H. Talbo, K. N. Truscott, D. A. Dougan, Modification of PATase by L/F-transferase generates a ClpS-dependent N-end rule substrate in Escherichia coli. *EMBO Journal* **28**, 1732-1744 (2009).
7. S. Gottesman, E. Roche, Y. Zhou, R. T. Sauer, The ClpXP and ClpAP proteases degrade proteins with carboxy-terminal peptide tails added by the SsrA-tagging system. *Genes & Development* **12**, 1338-1347 (1998).
8. S. K. Singh, R. Grimaud, J. R. Hoskins, S. Wickner, M. R. Maurizi, Unfolding and internalization of proteins by the ATP-dependent proteases ClpXP and ClpAP. *Proceedings of the National Academy of Sciences of the United States of America* **97**, 8898-8903. (2000).
9. J. M. Flynn *et al.*, Overlapping recognition determinants within the ssrA degradation tag allow modulation of proteolysis. *Proceedings of the National Academy of Sciences of the United States of America* **98**, 10584-10589. (2001).
10. L. M. Bittner, J. Arends, F. Narberhaus, Mini review: ATP-dependent proteases in bacteria. *Biopolymers* **105**, 505-517 (2016).
11. A. O. Olivares, T. A. Baker, R. T. Sauer, Mechanical protein unfolding and degradation. *Annual Review of Physiology* **80**, 413-429 (2018).

12. L. M. Bittner, J. Arends, F. Narberhaus, When, how and why? Regulated proteolysis by the essential FtsH protease in Escherichia coli. *Biological Chemistry* **398**, 625-635 (2017).
13. S. A. Mahmoud, P. Chien, Regulated proteolysis in bacteria. *Annual Review of Biochemistry* **87**, 677-696 (2018).
14. N. J. Kuhlmann, P. Chien, Selective adaptor dependent protein degradation in bacteria. *Current Opinion in Microbiology* **36**, 118-127 (2017).
15. M. Iconomou, D. N. Saunders, Systematic approaches to identify E3 ligase substrates. *Biochemical Journal* **473**, 4083-4101 (2016).
16. L. D. Rogers, C. M. Overall, Proteolytic post-translational modification of proteins: proteomic tools and methodology. *Molecular and Cellular Proteomics* **12**, 3532-3542 (2013).
17. N. J. Agard, J. A. Wells, Methods for the proteomic identification of protease substrates. *Current Opinion in Chemical Biology* **13**, 503-509 (2009).
18. M. Vizovisek, R. Vidmar, M. Fonovic, B. Turk, Current trends and challenges in proteomic identification of protease substrates. *Biochimie* **122**, 77-87 (2016).
19. P. F. Lange, C. M. Overall, Protein TAILS: when termini tell tales of proteolysis and function. *Current Opinion in Chemical Biology* **17**, 73-82 (2013).
20. C. Lopez-Otin, C. M. Overall, Protease degradomics: a new challenge for proteomics. *Nature Reviews: Molecular Cell Biology* **3**, 509-519 (2002).
21. O. Schilling, C. M. Overall, Proteomic discovery of protease substrates. *Current Opinion in Chemical Biology* **11**, 36-45 (2007).
22. F. Demir, S. Niedermaier, J. G. Villamor, P. F. Huesgen, Quantitative proteomics in plant protease substrate identification. *New Phytologist* **218**, 936-943 (2018).
23. L. Tsiatsiani *et al.*, The arabidopsis metacaspase9 degradome. *Plant Cell* **25**, 2831-2847 (2013).

24. K. Nishimura, Y. Kato, W. Sakamoto, Essentials of proteolytic machineries in chloroplasts. *Mol Plant* **10**, 4-19 (2017).
25. K. Nishimura, K. J. van Wijk, Organization, function and substrates of the essential Clp protease system in plastids. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **1847**, 915-930 (2015).
26. K. J. van Wijk, Protein maturation and proteolysis in plant plastids, mitochondria, and peroxisomes. *Annual Review of Plant Biology* **66**, 75-111 (2015).
27. Y. Kato, W. Sakamoto, FtsH protease in the thylakoid membrane: physiological functions and the regulation of protease activity. *Frontiers in Plant Science* **9**, 855 (2018).
28. A. Putarjunan, X. Liu, T. Nolan, F. Yu, S. Rodermel, Understanding chloroplast biogenesis using second-site suppressors of immutans and var2. *Photosynthesis Research* **116**, 437-453 (2013).
29. J. C. Moreno *et al.*, Temporal proteomics of inducible RNAi lines of Clp protease subunits identifies putative protease substrates. *Plant Physiology* **176**, 1485-1508 (2018).
30. M. L. Biniossek *et al.*, Identification of protease specificity by combining proteome-derived peptide libraries and quantitative proteomics. *Molecular and Cellular Proteomics* **15**, 2515-2524 (2016).
31. J. Tucher, A. Tholey, Multiplexed protease specificity profiling using isobaric labeling. *Methods in Molecular Biology* **1574**, 171-182 (2017).
32. J. Tucher *et al.*, LC-MS based cleavage site profiling of the proteases ADAM10 and ADAM17 using proteome-derived peptide libraries. *Journal of Proteome Research* **13**, 2205-2214 (2014).
33. U. Eckhard *et al.*, Active site specificity profiling of the matrix metalloproteinase family: proteomic identification of 4300 cleavage sites by nine MMPs explored with structural and synthetic peptide cleavage analyses. *Matrix Biology* **49**, 37-60 (2016).

34. O. Barre *et al.*, Cleavage specificity analysis of six type II transmembrane serine proteases (TTSPs) using PICS with proteome-derived peptide libraries. *PloS One* **9**, e105984 (2014).

35. M. Paireder *et al.*, The papain-like cysteine proteinases NbCysP6 and NbCysP7 are highly processive enzymes with substrate specificities complementary to Nicotiana benthamiana cathepsin B. *Biochim Biophys Acta Proteins Proteom* **1865**, 444-452 (2017).

36. J. M. Flynn, S. B. Neher, Y. I. Kim, R. T. Sauer, T. A. Baker, Proteomic discovery of cellular substrates of the ClpXP protease reveals five classes of ClpX-recognition signals. *Molecular Cell* **11**, 671-683. (2003).

37. J. M. Flynn, I. Levchenko, R. T. Sauer, T. A. Baker, Modulating substrate choice: the SspB adaptor delivers a regulator of the extracytoplasmic-stress response to the AAA+ protease ClpXP for degradation. *Genes & Development* **18**, 2292-2301 (2004).

38. S. B. Neher *et al.*, Proteomic profiling of ClpXP substrates after DNA damage reveals extensive instability within SOS regulon. *Molecular Cell* **22**, 193-204 (2006).

39. D. B. Trentini *et al.*, Arginine phosphorylation marks proteins for degradation by a Clp protease. *Nature* **539**, 48-+ (2016).

40. J. Arends, N. Thomanek, K. Kuhlmann, K. Marcus, F. Narberhaus, In vivo trapping of FtsH substrates by label-free quantitative proteomics. *Proteomics* **16**, 3161-3172 (2016).

41. K. Westphal, S. Langklotz, N. Thomanek, F. Narberhaus, A trapping approach reveals novel substrates and physiological functions of the essential protease FtsH in Escherichia coli. *Journal of Biological Chemistry* **287**, 42962-42971 (2012).

42. M. Opalinska, K. Parys, H. Janska, Identification of physiological substrates and binding partners of the plant mitochondrial protease FTSH4 by the trapping approach. *International Journal of Molecular Sciences* **18**, (2017).

43. J. Arends *et al.*, An integrated proteomic approach uncovers novel substrates and functions of the Lon protease in Escherichia coli. *Proteomics* **18**, e1800080 (2018).
44. S. M. Doyle, S. Wickner, Hsp104 and ClpB: protein disaggregating machines. *Trends in Biochemical Sciences* **34**, 40-48 (2009).
45. T. Haslberger, B. Bukau, A. Mogk, Towards a unifying mechanism for ClpB/Hsp104-mediated protein disaggregation and prion propagation. *Biochemistry and Cell Biology* **88**, 63-75 (2010).
46. R. C. Mishra, A. Grover, ClpB/Hsp100 proteins and heat stress tolerance in plants. *Critical Reviews in Biotechnology* **36**, 862-874 (2016).
47. J. Weibezahn, C. Schlieker, B. Bukau, A. Mogk, Characterization of a trap mutant of the AAA+ chaperone ClpB. *Journal of Biological Chemistry* **278**, 32608-32617 (2003).
48. J. Krajewska, Z. Arent, M. Zolkiewski, S. Kedzierska-Mieszkowska, Isolation and identification of putative protein substrates of the AAA+ molecular chaperone ClpB from the pathogenic spirochaete Leptospira interrogans. *International Journal of Molecular Sciences* **19**, (2018).
49. J. W. Graham, M. G. Lei, C. Y. Lee, Trapping and identification of cellular substrates of the Staphylococcus aureus ClpC chaperone. *Journal of Bacteriology* **195**, 4506-4516 (2013).
50. C. Montandon, G. Friso, J. Liao, J. Choi, K.J. van Wijk, Chloroplast CLPC1 trapping *in vivo*; recruitment of the CLPPR core complex and CLP candidate substrates. *in revision*, (2019).
51. J. Feng *et al.*, Trapping and proteomic identification of cellular substrates of the ClpP protease in Staphylococcus aureus. *Journal of Proteome Research* **12**, 547-558 (2013).
52. N. H. Bhat, R. H. Vass, P. R. Stoddard, D. K. Shin, P. Chien, Identification of ClpP substrates in Caulobacter crescentus reveals a role for regulated proteolysis in bacterial development. *Molecular Microbiology* **88**, 1083-1092 (2013).

53. J.-Y. R. Liao, Friso, G., Kim,J., van Wijk, K.J., Consequences of the loss of catalytic triads in chloroplast CLPPR protease core complexes in vivo. *PlantDirect* **in press**, (2018).

54. M. R. Maurizi, W. P. Clark, S. H. Kim, S. Gottesman, ClpP represents a unique family of serine proteases. *Journal of Biological Chemistry* **265**, 12546-12552 (1990).

55. I. Wagner, L. van Dyck, A. S. Savel'ev, W. Neupert, T. Langer, Autocatalytic processing of the ATP-dependent PIM1 protease: crucial function of a pro-region for sorting to mitochondria. *EMBO Journal* **16**, 7317-7325 (1997).

56. M. Koppen, F. Bonn, S. Ehses, T. Langer, Autocatalytic processing of m-AAA protease subunits in mitochondria. *Molecular Biology of the Cell* **20**, 4216-4224 (2009).

57. D. Mellacheruvu *et al.*, The CRAPome: a contaminant repository for affinity purification-mass spectrometry data. *Nature Methods* **10**, 730-736 (2013).

58. M. A. Humbard, S. Surkov, G. M. De Donatis, L. M. Jenkins, M. R. Maurizi, The N-degradome of Escherichia coli: limited proteolysis in vivo generates a large pool of proteins bearing N-degrons. *Journal of Biological Chemistry* **288**, 28913-28924 (2013).

59. K. Nishimura *et al.*, ClpS1 is a conserved substrate selector for the chloroplast Clp protease system in Arabidopsis. *Plant Cell* **25**, 2276-2301 (2013).

60. L. L. Sjogren, T. M. MacDonald, S. Sutinen, A. K. Clarke, Inactivation of the clpC1 gene encoding a chloroplast Hsp100 molecular chaperone causes growth retardation, leaf chlorosis, lower photosynthetic activity, and a specific reduction in photosystem content. *Plant Physiology* **136**, 4114-4126 (2004).

61. D. Constan, J. E. Froehlich, S. Rangarajan, K. Keegstra, A stromal Hsp100 protein is required for normal chloroplast development and function in Arabidopsis. *Plant Physiology* **136**, 3605-3615 (2004).

62. K. Nishimura *et al.*, Discovery of a unique Clp component, ClpF, in chloroplasts: a proposed binary ClpF-ClpS1 adaptor complex functions in substrate recognition and delivery. *Plant Cell* **27**, 2677-2691 (2015).

63. J. Apitz *et al.*, Posttranslational control of ALA synthesis includes GluTR degradation by Clp protease and stabilization by GluTR-binding protein. *Plant Physiology* **170**, 2040-2051 (2016).

64. A. Bensimon, A. J. Heck, R. Aebersold, Mass spectrometry-based proteomics and network biology. *Annual Review of Biochemistry* **81**, 379-405 (2012).

65. R. Aebersold, M. Mann, Mass-spectrometric exploration of proteome structure and function. *Nature* **537**, 347-355 (2016).

66. P. Lossl, M. van de Waterbeemd, A. J. Heck, The diverse and expanding role of mass spectrometry in structural and molecular biology. *EMBO Journal* **35**, 2634-2657 (2016).

CHAPTER THREE

CONSEQUENCES OF THE LOSS OF CATALYTIC TRIADS IN CHLOROPLAST CLPPR PROTEASE CORE COMPLEXES *IN VIVO*[1,2]

3.1 ABSTRACT

The essential chloroplast CLP protease system consists of a tetradecameric proteolytic core with catalytic P (P1, 3-6) and non-catalytic R (R1-4) subunits, CLP chaperones and adaptors. The chloroplast CLP complex has a total of ten catalytic sites, but is not known how many of these catalytic sites can be inactivated before plants lose viability. Here we show that CLPP3 and the catalytically inactive variant CLPP3S164A fully complements the developmental arrest of the *clpp3-1* null mutant, even under environmental stress. In contrast, whereas the inactive variant CLPP5S193A assembled into the CLP core, it cannot rescue the embryo lethal phenotype of the *clpp5-1* null mutant. This shows that CLPP3 makes a unique structural contribution but its catalytic site is dispensable, whereas the catalytic activity of CLPP5 is essential. Mass spectrometry of affinity-purified CLP cores of the complemented lines showed highly enriched CLP cores. Other chloroplast proteins were co-purified with the CLP cores and candidate substrates. A strong overlap of co-purified proteins between the CLP core complexes with active and inactive subunits indicates that CLP cores with reduced number of catalytic sites do not over-accumulate substrates, suggesting that the bottleneck for degradation is likely substrate

recognition and unfolding by CLP adaptors and chaperones, upstream of the CLP core.

---

[1] Previously published in Liao et al, Plant Direct 2018 2 (10) e00086. Copyright owned by the American Society of Plant Biologists (ASPB).

[2]Jui-Yun Rei Liao conducted all the plant work (with the aid from Dr. Jitae Kim for site-directed mutagenesis) and prepared for the protein samples for MS analysis. Dr. Giulia Friso conducted all MS analysis.

3.2 INTRODUCTION

ATP-dependent Clp proteases are present in bacteria, as well as mitochondria and plastids (organelles of bacterial origin), where they regulate accumulation levels of a broad range of substrates (*1-5*). The first step in the CLP degradation process requires the recognition of substrates by the CLP AAA+ chaperones, possibly aided by specific adaptors (also named recognins) that recognize and deliver specific substrates (*6, 7*). The ATP-dependent CLP chaperones then dock onto CLP protease core complexes consisting of two stacked heptameric rings, and unfold and direct substrates into the CLP protease complex (*8*). The substrates are cleaved within the CLP protease complex and short (~7-10 amino acids) peptide fragments are released through lateral pores in the CLP protease (*2, 3, 9*). To ensure optimal cellular levels of functional proteins and to remove unwanted proteins while avoiding uncontrolled degradation, substrate recognition and delivery by the adaptors and chaperones must be tightly regulated and depends also on the availability and exposure of degrons in the substrates.

The CLP proteolytic core *in E. coli* consists of 14 identical CLPP subunits that belong to the family of serine-type proteases (*10*). The photosynthetic bacterium *Synechococcus elongatus* contains three CLPP genes and one CLPR gene (*11*), which is structurally similar to CLPP but lacks the catalytic residues for peptide bond hydrolysis. These cyanobacterial proteins assemble into an essential CLPR/P3 complex (identical heptameric rings with P3:R in a 3:4 ratio) and a non-essential CLPP1/P2 complex (*12, 13*). The apicoplast of *Plasmodium falsiparum* accumulates both a single CLPP and CLPR protein but they do appear to form separate homo-oligomers (*14*). Compared to these homologs, the chloroplast CLP system is far more diversified and complex (*4, 15, 16*).

The chloroplast CLP protease complex in Arabidopsis contains five different CLPP subunits namely plastid-encoded CLPP1 and the nuclear-encoded CLPP3-6, and four different non-catalytic CLPR subunits CLPR1-4. The five CLPP subunits with conserved catalytic sites accumulate in a 3:1:2:3:1 ratio for P1:P3:P4:P5:P6, whereas the non-catalytic CLPR subunits are each present in one copy. Three copies of CLPP1 and the four CLPR subunits form the heptameric R-ring, whereas CLPP3-6 forms the heptameric P-ring. Therefore the chloroplast CLP complex has a total of ten catalytic sites, of which seven are in the P-ring and only three are in the R-ring. It is not known how many and which of these catalytic sites can be inactivated before plants lose viability. The chloroplast CLP core also associates with two plant specific proteins ClpT1 and ClpT2. These CLPT subunits likely function in CLPPR core formation, stabilization and activation (*17, 18*).

Arabidopsis CLPP and CLPR proteins are not very conserved to each other at the primary sequence level (24-48% identities between CLPPs; 28-38% identities between CLPRs). Severe phenotypes were found in Arabidopsis *clpp* and *clpr* mutants and in tobacco RNAi lines (*19*) at various developmental stages (*20-24*). For example, null mutants in *CLPP3* (*clpp3-1*) are seedling lethal and are arrested in the cotyledon stage. However, the addition of sucrose breaks this developmental arrest and plants form pale-green leaves and eventually flower and produce seed (*20*). In contrast, null mutants in *CLPP5* (*clpp5-1*) are embryo lethal (*21*). These phenotypes showed that both CLP3 and CLP5 are required for plant development, and we speculated that the differential phenotype between *clpp3-1* and *clpp5-1* is in part due to the higher copy number of CLPP5 (three per complex) than CLPP3 (one per complex). Through site-directed mutagenesis and complementation of null mutants, this study tests if the catalytic contribution of CLPP3 and CLPP5 can be inactivated (while keeping the structural contribution) without functional consequences for growth and development.

So far, several (candidate) chloroplast CLP substrates have been identified based on their direct interaction with the CLPS1 adaptor, including GLUTR (*25*), which was subsequently confirmed as a CLP substrate (*26*). Other candidate CLPS1 substrates are four enzymes in the chloroplast shikimate pathway (*25*). Furthermore, systematic screening of protein stability of the thylakoid copper transporter, PAA2/HMA8 (P-type ATPase of Arabidopsis2/Heavy-metal-associated8) in various plastid protease mutants identified the CLPPR core and CLPC1 as essential components to degrade PAA2 under copper replete conditions (*27*). Additionally, DEOXYXYLULOSE 5-PHOSHATE (DXS) in the chloroplast isoprenoid

biosynthesis pathway and PHYTOENE SYNTHASE (PSY) involved in carotenoid biosynthesis were reported as a CLP substrates (*28, 29*). Other candidate substrates have been suggested based on comparative proteome analysis of a range of *clp* mutants in Arabidopsis, rice and tobacco, but here it is hard to distinguish between direct and indirect effects - reviewed in (*30*) and (*31, 32*).

An alternative approach to identify CLP substrates, termed substrate trapping, has been reported for various bacterial and fungal CLP protease systems (*33-35*). In this approach, the catalytic activity of the CLP core is inactivated through site-mutagenesis of the serine residue within the catalytic triad (Ser-His-Asp), resulting in accumulation of substrates within the CLP core central cavity and facilitating their identification by tandem mass spectrometry (MS/MS). This *in vivo* trapping strategy has been applied for the homo-tetrameric CLP complexes in the gram-negative bacterium *E. coli* (*33*), the gram-positive *Staphylococcus aureus* (*34*) and *Bacillus subtilis* (*35*), as well as the fungus *Podospora anserine* (*36*), but not yet in plants. It is important to note that the inactivation of the CLP system in these species does not greatly affect viability. The subunit complexity and the essential nature of the chloroplast CLP core makes an *in vivo* CLP protease trapping approach in plants more challenging than in bacteria and fungi. In this study we used this approach for the *Arabidopsis* chloroplast CLP protease, generating *in vivo* tagged CLP core complexes containing inactive serine to alanine variants in either CLPP3 or CLPP5. MS/MS of affinity-purified tagged CLP cores from the various transformants showed that we successfully obtained highly enriched STREPII-tagged CLP cores with all catalytically active CLPP3 and CLPP5 copies and with catalytically inactive CLPP3 or

CLPP5 copies. Other chloroplast proteins were co-purified with the CLP cores and we evaluate and discuss their significance as candidate substrates in the context of bottlenecks in the chloroplast CLP adaptor-chaperone-protease system.

3.3 RESULTS AND DISCUSSION

3.3.1 Phenotypic analysis shows that the catalytic activity of CLPP3 is not needed for function

The catalytic triad Ser-His-Asp in CLPP proteases is widely conserved across bacterial CLPP proteins as well mitochondrial and plastid CLPP (but not in the catalytically inactive CLPR proteins) (*9, 15, 37*) (Supplemental Figure 1). Changing the serine residue into an alanine is sufficient to completely block proteolytic activity by CLPP proteases in all known cases, *e.g. E. coli* (*38*)*, Mycobacterium tuberculosis* (*39*), *Bacillus subtilis* (*35*) and *Synechococcus elongates* (*13*).

To determine if the contribution of CLPP3 to the core complex is through its unique structure or also by contributing a catalytic site to the core complex, we transformed heterozygous *clpp3-1* mutants with genomic *CLPP3* with a C-terminal STREPII tag (*CLPP3-STREPII*), or the same genomic construct but with a serine to alanine mutation in the catalytic site (*CLPP3S164A-STREPII*), rendering CLPP3 catalytically inactive. Following growth on selective medium and genotyping of the T1 and T2 populations, we obtained multiple homozygous *clpp3-1* lines expressing either *CLPP3-STREPII* or *CLPP3S164A-STREPII* (Figure 1A,B; Supplemental Figure 2A,B,C). No significant visible differences were observed among these complemented *clpp3-1* lines and wt, as they all germinated and developed without the need for

sucrose, as well as displayed normal green rosettes and similar growth and development (Figure 1A,B). In contrast, *clpp3-1* nulls were arrested at the cotyledon stage (Figure 1A), as described previously (*20*).

The point mutation in CLPP3 was confirmed by DNA sequencing of RT-PCR products generated by *CLPP3* and *STREPII* specific primers (Figure 1C). Furthermore, MS/MS of affinity purified CLP complexes identified the specific point mutation S164A in CLPP3 (Figure 1D; Supplemental Figure 3), thus further confirming that the point mutation in CLPP3 was successfully introduced into these lines. Accumulation of the transgenic CLPP3-STREPII and CLPP3S164A-STREPII proteins *in vivo* was also confirmed by immunoblotting of the denaturing soluble proteome separated by SDS-PAGE and detected with anti-STREPII and anti-CLPP3 antisera (Figure 1E). To test if STREPII-tagged CLPP3 and CLPP3S164A proteins normally assembled into CLPRT complexes, soluble leaf proteomes were separated by native gel electrophoresis (BN-PAGE), followed by immunoblotting (Figure 1F). The similar size and migration pattern of CLPPRT complexes on BN-PAGE for wt and complemented lines indicated that a) these STREPII-tagged transgenic proteins assembled into CLPPRT complexes *in vivo,* and b) the S-to-A change in CLPP3 and the STREPII tag did not interfere with the assembly state of CLPPRT complexes. Collectively, this shows that both catalytically active CLPP3-STREPII and catalytic inactive CLPP3S164A-STREPII successfully complement the developmental arrest of *clpp3-1* null mutants. It can thus be concluded that CLPP3 makes a unique structural contribution to the CLPPR complex but that its catalytic activity is dispensable for plant growth, development and stress responses.

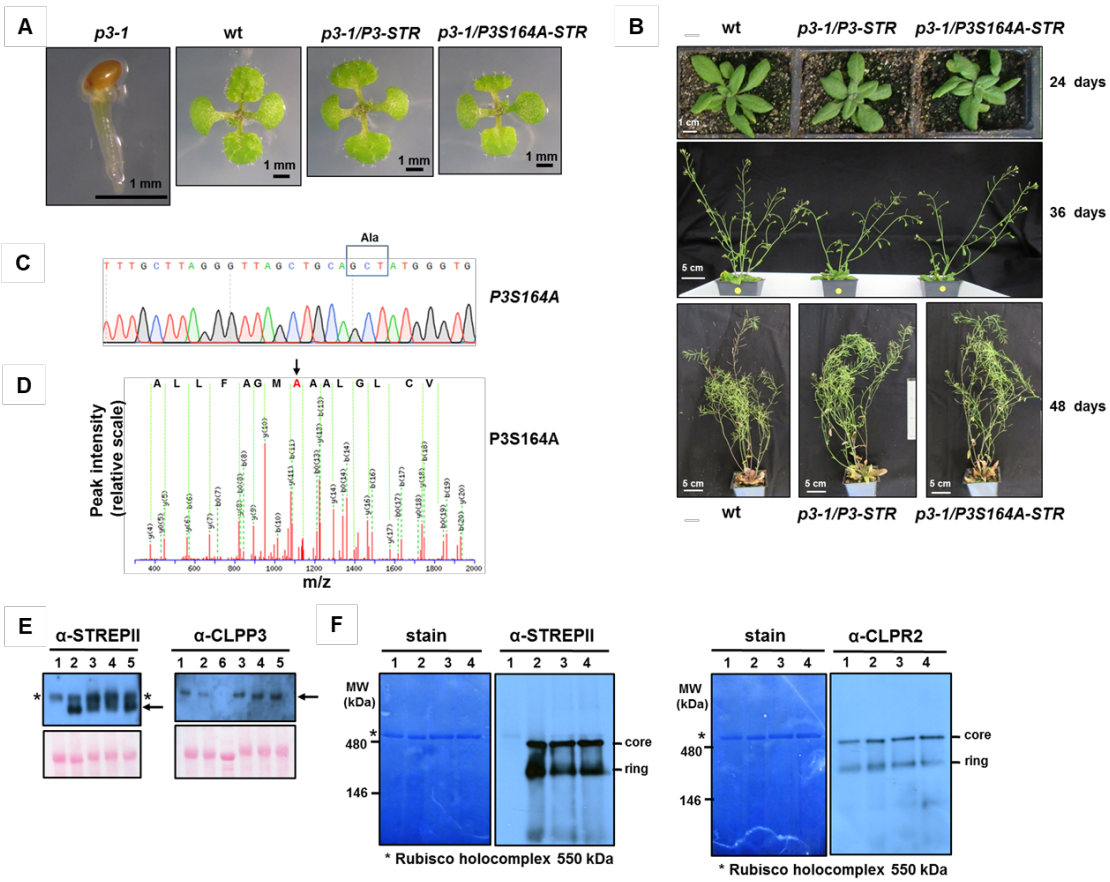
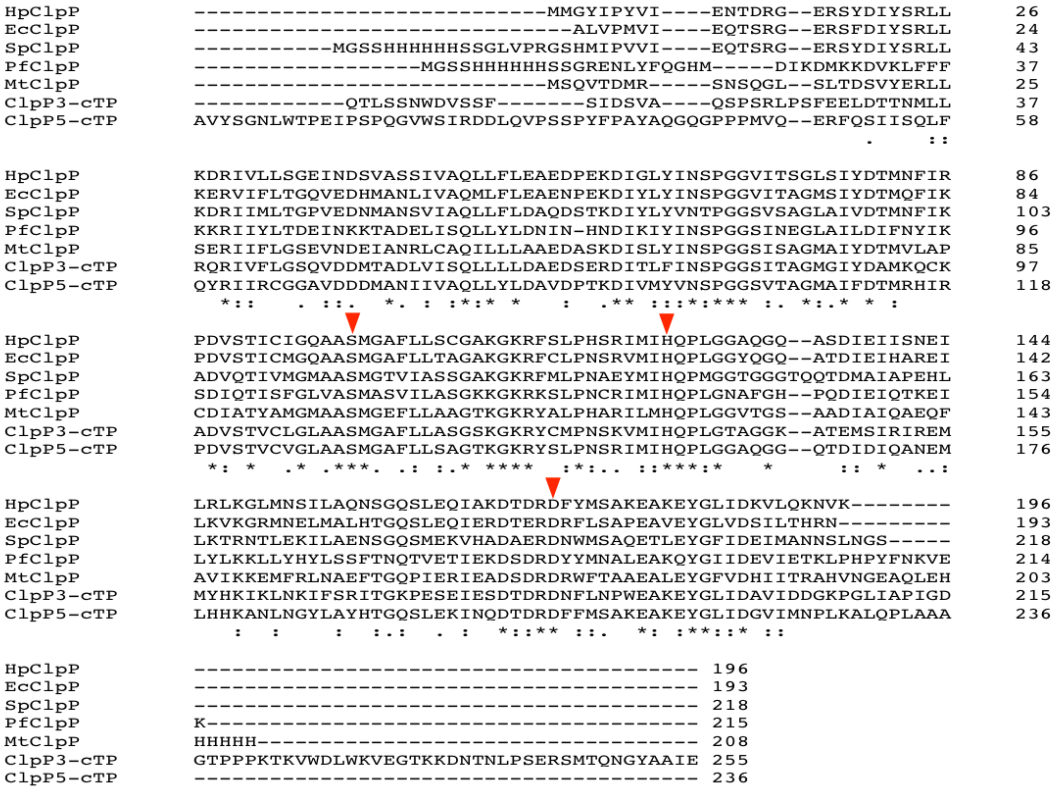


Figure 1. **Characterization of *clpp3-1* complemented with *CLPP3-STREPII* and *CLPP3S164A-STREPII.***

A, Seedlings of wt, *clpp3-1* and *clpp3-1* complemented with either *CLPP3-STREPII* or *CLPP3S164A-STREPII* grown on half MS medium (no sucrose) for 12 days under 16/8 hr light/dark cycle at 70 µmol photons.m^{-2}.s^{-1}. Plants are homozygous for the STREPII transgenes. **B,** Different developmental stages of soil-grown wt and *clpp3-1* complemented with either *CLPP3-STREPII* or *CLPP3S164A-STREPII*. Plants were grown on soil under 16/8 hr light/dark cycle at 120-150 µmol photons.m^{-2}.s^{-1}. Plants are homozygous for the STREPII transgenes. **C,** Confirmation of the point mutation in the catalytic site of *CLPP3S164A-STREPII* by DNA sequencing of RT-PCR products (primers #6 and #9) amplified from transgenic plants. The codon responsible for the serine to alanine point mutations is boxed and is TCT in wt but GCT in the mutant. **D,** Confirmation of the point mutation S164 to A164 in the catalytic site of CLPP3S164A-STREPII by MS/MS of the tryptic peptide

(ADVSTVCLGLAAA$_{164}$MGAFLLASGSK) generated by tryptic digestion of affinity purified CLP complexes. The MS/MS spectrum is from a doubly charged precursor ion with m/z of 1155.5967 (2^+) with MASCOT ion score of 131 (0.38 ppm error) and supports the residue A164. The partial peptide sequence listed above the spectrum (ALLFAGMA$_{164}$AALGLCV) shown is based on y-ions explaining the reverse order of amino acids. A list of b- and y-ions is provided in Supplemental Figure 3A. An example of an MS/MS spectrum of the analogous CLPP3 wild-type peptide is provided in Supplemental Figure 3B. **E,** SDS-PAGE gel and immunoblotting of total soluble protein from wt (lane 1), *clpp3-1* with CLPP3**-STREPll (lane 2), total soluble protein from *clpp3-1* with *CLPP3-STREPII* (lane 3) and *clpp3-1* with *CLPP3S164A-STREPII* (lanes 4 and 5), and stromal protein from *clpp3-1* (lane 6). Arrows indicate CLPP3-STREPll or CLPP3S164A-STREPll proteins. * indicates the nonspecific reaction. The Ponceau red stain of the blot illustrates protein loading. **35-S driven cDNA of CLPP3-STREPII as described in (*20*); all other lines and samples in the current study are using genomic CLP DNA for transgene expression (see methods). Anti-CLPP3 or anti-STREPII serum was used. **F,** Native gels and immunoblotting of stromal proteomes of wt (lane 1), *clpp3-1* with CLPP3**-STREPll (lane 2), *clpp3-1* with CLPP3-STREPll ((lane 3), and *clpp3-1* with CLPP3S164A-STREPll (lane 4). The intact CLPPR core and individual rings are indicated. * indicates the Rubisco complex. Anti-STREPII or anti-CLPR2 serum was used for visualization of the core CLP complex and R-ring. **35-S driven cDNA of CLPP3-strepll as described in (*20*).

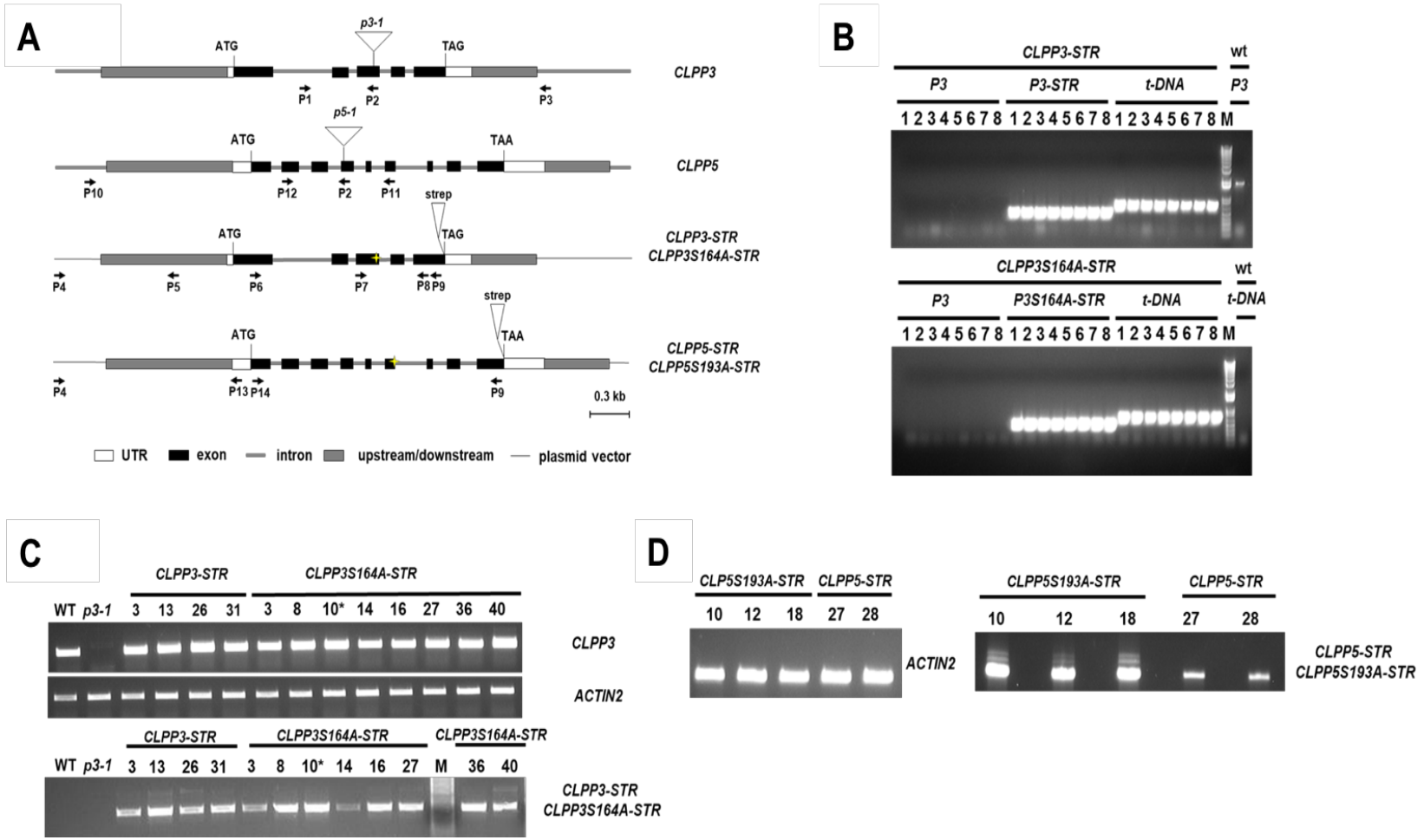


Supplemental Figure 1. Sequence alignment of CLPP3 and CLPP5 with CLPP homologs.

Crystalized CLPP homologs from *Helicobacter pylori* (hp), *Escherichia coli* (Ec), *Streptococcus pnneumoniae* (Sp), *Plasmodium falciparum* (Pf), and *Mycobacterium tuberculosis* (Mt) are aligned and compared with CLPP3 and CLPP5 without cTP. The residues in the catalytic triad are indicated with arrowheads. sequence identity (*); high sequence similarity (:); low level of similarity (.)

Supplemental Figure 2. Genotyping and RT-PCR analysis of transgenic lines.

A, Gene models and the transgenic constructs for complementation. Primers used for genotyping and RT-PCR analysis are numbered (see primer # in Table S1). The star highlights the Ser-to-Ala change. The upstream/downstream region and the plastid vector shown are not to scale. **B,** Genotyping analysis of the endogenous *CLPP3* (P1+P3), transgenic *CLPP3-STREPII* or *CLPP3S164A-STREPII* (P4+P5), or *clpp3-1* mutant allele (P1+P2). Eight individual T2 plants from one *CLPP3-STREPII* T1 line (upper panel) or *CLPP3S164A-STREPII* T1 line (lower panel) were analyzed. wt was for the positive control *CLPP3* and the negative control for *p3-1* null. *CLPP3-STR* T2 or *CLPP3S164A-STR* T2 indicates the genomic DNA extracted from the which T2 lines. *P3*, *P3-STR*, *P3S164A-STR*, and *t-DNA* indicates the primer pairs used to specifically amplify genomic *CLPP3*, transgenic *CLPP3-STREPII*, transgenic *CLPP3S164A-STREPII*, and t-DNA insertion, respectively. **C,** RT-PCR analysis of *CLPP3-STREPII* or *CLPP3S164A-STREPll* expression (P7+P8 or P6+P9) in the transgenic T1 lines. Four putative *CLPP3-STREPII* T1 or eight putative *CLPP3S164A-STREPll* T1 plants were analyzed using either *ACTIN2* was used as an internal control. * - removed from further analysis. *CLPP3-STR* T1 or *CLPP3S164A-STR* T1 indicates the RNA extracted from the which T1 lines. *CLPP3*, *CLPP3-STR*, *CLPP3S164A-STR* indicates the primer pairs used to specifically amplify *CLPP3*, *CLPP3-STREPII*, and *CLPP3S164A-STREPII* fragment, respectively. **D,** RT-PCR analysis of *CLPP5-STREPII* or *CLPP5S193A-STREPll* expression (P14+P9) in the two putative *CLPP5-STR* and three putative *CLPP5S193A-STR* T1 lines. *ACTIN2* was used as an internal control. *CLPP5-STR* T1 or *CLPP5S193A-STR* T1 indicates the RNA extracted from the which T1 lines. *CLPP5-STR*, *CLPP5S193A-STR* indicates the primer pairs used to specifically amplify *CLPP5-STREPII*, and *CLPP5S193A-STREPII* fragment, respectively.


A
ATG
p3-1
TAG
CLPP3
P1
P2
P3
p5-1
ATG
TAA
CLPP5
P10
P12
P2
P11
strep
ATG
TAG
CLPP3-STR
CLPP3S164A-STR
P4
P5
P6
P7
P8 P9
strep
ATG
TAA
CLPP5-STR
CLPP5S193A-STR
P4
P13 P14
P9
0.3 kb
UTR
exon
intron
upstream/downstream
plasmid vector
B
CLPP3-STR
wt
P3
P3-STR
t-DNA
P3
1 2 3 4 5 6 7 8 1 2 3 4 5 6 7 8 1 2 3 4 5 6 7 8 M
CLPP3S164A-STR
wt
P3
P3S164A-STR
t-DNA
t-DNA
1 2 3 4 5 6 7 8 1 2 3 4 5 6 7 8 1 2 3 4 5 6 7 8 M
C
CLPP3-STR
CLPP3S164A-STR
WT p3-1 3 13 26 31 3 8 10* 14 16 27 36 40
CLPP3
ACTIN2
CLPP3-STR
CLPP3S164A-STR
CLPP3S164A-STR
WT p3-1 3 13 26 31 3 8 10* 14 16 27 M 36 40
CLPP3-STR
CLPP3S164A-STR
D
CLP5S193A-STR
CLPP5-STR
10 12 18 27 28
ACTIN2
CLPP5S193A-STR
CLPP5-STR
10 12 18 27 28
CLPP5-STR
CLPP5S193A-STR

Supplemental Table 1. Primer sets used in P3-P5 study			
primer sets for cloning	**Function**	**primer sequence**	**primer #**
primer set 1	gP3_N_cloning	CAC CAA GCA GCT AAA GCA CAT AAG TTT	
	gP3_C_cloning	CGG CTC TAT GTC CGT CGT TTC	
	gP5_N_cloning	CAC CTA CGA GAA AGC TAA CGC CAC G	
	gP5_C_cloning	ACG GGA CGG GTT TGA CAG GAC GTT	
primer set 2	P3_strepII_R	CTA CTT CTC GAA TTG AGG ATG AGA CCA TTC AAT GGC GGC ATA ACC ATT CTG TGT CAT GG	
	P3_strepII_F	TGG TCT CAT CCT CAA TTC GAG AAG TAG AAC TGT TGT TGC AGC GTT TAC GCC TTT TAT AT	
	P5_strepII_F	TTA CTT CTC GAA TTG AGG ATG AGA CCA AGC TGC TGC AAG TGG CTG GAG AGC TTT AAG AG	
	P5_strepII_R	TGG TCT CAT CCT CAA TTC GAG AAG TAA TCG CCT AAA GGT AGT GGT TCA GCT TTA GCA CT	
primer set 3	P3_StoA_F	CTT AGG GTT AGC TGC AGC TAT GGG TGC GTT TC	
	P3_StoA_R	GAA ACG CAC CCA TAG CTG CAG CTA ACC CTA AG	
	P5_StoA_F	TGT TGG TCT AGC TGC TGC GTA AAA CTT TCA CC	
	P5_StoA_R	GGT GAA AGT TTT ACG CAG CAG CTA GAC CAA CA	
primer sets for genotype			
primer set 4	P3_wtspecific_F	GTT TGA TGA GAG TCA GAG AGA GTA GAG A	1
	P3_wtspecific_R	CAG ATT CAT CGG GAG AAG GAG GAA G	3
	P5_wtspecific_F	CCC ATC GAC GTC CTT CGA ATG G	10
	P5_wtspecific_R	CAT CAG GCC GGA TGT GCC TCA TAG TAT C	11
primer set 5	P3transgene_specific_F (vector)	CTC GAG GCG CGC CAA GCT ATC	4
	P3transgene_specific_R	GCA CAT CAA CAG TCT TGC TTG ACA C	5
	P5transgene_specific_F (vector)	CTC GAG GCG CGC CAA GCT ATC	4
	P5transgene_specific_R	GAG CCA TTT TCT TGT TAT CAA TTT GGG GAT C	13
primer set 6	P3tDNA_specific_F	GTT TGA TGA GAG TCA GAG AGA GTA GAG A	1
	P3tDNA_specific_R (LBb1.3)	ATT TTG CCG ATT TCG GAA C	2
	P5tDNA_specific_F	CGT CTC CTC AAG GAG TTT GGT CCA TTA G	12
	P5tDNA_specific_R (LBb1.3)	ATT TTG CCG ATT TCG GAA C	2
primer sets for RT-PCR			
primer set 7	P3_RT_F1	CTA GTA ACT GGG ATG TAT CTA GCT TCT CCA TTG	6
	P3_RT_R1(strep)	TCA CTT CTC GAA TTG AGG ATG AGA CCA	9
	P3_RT_F2	GCT TAG GGT TAG CTG CAT CTA TGG GTG	7
	P3_RT_R2	CAA TGG CGG CAT AAC CAT TCT GTG TC	8
primer set 8	P5_RT_F	ATG GCT CAT GCT TGC GTC TC	14
	P5_RT_R (strep)	TCA CTT CTC GAA TTG AGG ATG AGA CCA	9
primer set 9	Actin2_F	CAA ACG AGG GCT GGA ACA AGA CT	
	Actin2_R	GCA ACT GGG ATG ATA TGG AAA AGA	

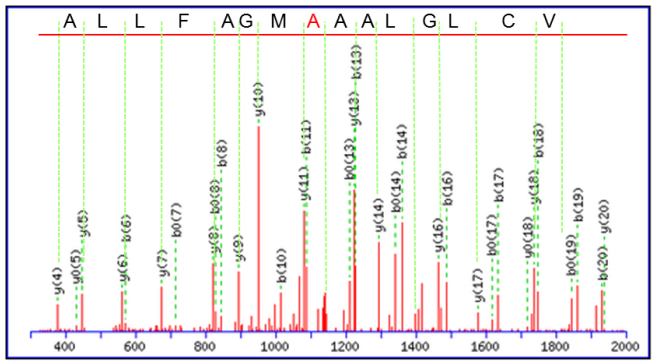
Supplemental Figure 3. **MS/MS-based verification of the S164A mutation in CLPP3S164A-STREPII and comparison to CLPP3-STREPII.**

A, Identification of the point mutation S164 to A164 in the catalytic site of CLPP3S164A-STREPII by MS/MS of the tryptic peptide (ADVSTVCLGLAAAMGAFLLASGSK) generated by tryptic digestion of affinity purified CLP complexes. The MS/MS spectrum is from a doubly charged precursor ion with m/z of 1155.5967 (2^+) with MASCOT ion score of 131 (0.38 ppm error) and supports the residue A164. The partial peptide sequence listed above the spectrum (ALLFAGMA_{164}AALGLCV) shown is based on y-ions explaining the reverse order of amino acids. A full list of b- and y-ions is listed. **B,** An example of an MS/MS spectrum of the CLPP3 wild-type peptide covering the region around S164. MS/MS of the tryptic peptide (ADVSTVCLGLAASMGAFLLASGSK) generated by tryptic digestion of affinity purified CLP complexes. The MS/MS spectrum is from a doubly charged precursor ion with m/z of 1163.5893 (2^+) with MASCOT ion score of 131 (-2.92 ppm error) and supports the residue S164. The partial peptide sequence listed above the spectrum (ALLFAGMS_{164}AALGLCV) shown is based on y-ions explaining this reads in reverse order. A full list of b- and y-ions is listed.

A

Precursor Ion = 1155.5967 (2+)
Mascot Score = 131
Measured Mass = 2309.1768
Calculated Mass = 2309.1759
ppm = 0.38

CLPP3S164A-STREPII



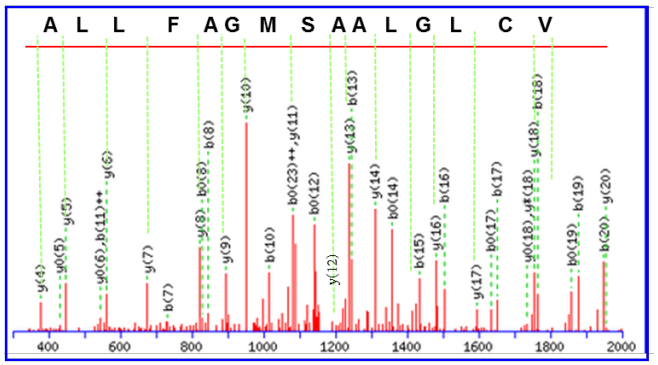
ADVSTVCLGLAAA$_{164}$MGAFLLASGSK

#	b	b++	b0	b0++	Seq.	y	y++	y*	y*++	y0	y0++	#
1	72.0444	36.5258			A							24
2	187.0713	94.0393	169.0608	85.0340	D	2239.1461	1120.0767	2222.1196	1111.5634	2221.1355	1111.0714	23
3	286.1397	143.5735	268.1292	134.5682	V	2124.1192	1062.5632	2107.0926	1054.0499	2106.1086	1053.5579	22
4	373.1718	187.0895	355.1612	178.0842	S	2025.0507	1013.0290	2008.0242	1004.5157	2007.0402	1004.0237	21
5	474.2195	237.6134	456.2089	228.6081	T	*1938.0187*	969.5130	1920.9922	960.9997	1920.0082	960.5077	20
6	573.2879	287.1476	555.2773	278.1423	V	1836.9710	918.9892	1819.9445	910.4759	1818.9605	909.9839	19
7	733.3185	367.1629	715.3080	358.1576	C	*1737.9026*	869.4550	1720.8761	860.9417	1719.8921	860.4497	18
8	846.4026	423.7049	828.3920	414.6996	L	*1577.8720*	789.4396	1560.8454	780.9264	1559.8614	780.4343	17
9	903.4240	452.2157	885.4135	443.2104	G	*1464.7879*	732.8976	1447.7614	724.3843	1446.7773	723.8923	16
10	1016.5081	508.7577	998.4975	499.7524	L	1407.7664	704.3869	1390.7399	695.8736	1389.7559	695.3816	15
11	1087.5452	544.2762	1069.5347	535.2710	A	*1294.6824*	647.8448	1277.6558	639.3316	1276.6718	638.8395	14
12	1158.5823	579.7948	1140.5718	570.7895	A	*1223.6453*	612.3263	1206.6187	603.8130	1205.6347	603.3210	13
13	1229.6195	615.3134	1211.6089	606.3081	A	1152.6082	576.8077	1135.5816	568.2944	1134.5976	567.8024	12
14	1360.6599	680.8336	1342.6494	671.8283	M	*1081.5710*	541.2892	1064.5445	532.7759	1063.5605	532.2839	11
15	1417.6814	709.3443	1399.6708	700.3391	G	*950.5306*	475.7689	933.5040	467.2556	932.5200	466.7636	10
16	1488.7185	744.8629	1470.7079	735.8576	A	*893.5091*	447.2582	876.4825	438.7449	875.4985	438.2529	9
17	1635.7869	818.3971	1617.7764	809.3918	F	*822.4720*	411.7396	805.4454	403.2264	804.4614	402.7343	8
18	1748.8710	874.9391	1730.8604	865.9339	L	*675.4036*	338.2054	658.3770	329.6921	657.3930	329.2001	7
19	1861.9551	931.4812	1843.9445	922.4759	L	*562.3195*	281.6634	545.2930	273.1501	544.3089	272.6581	6
20	1932.9922	966.9997	1914.9816	957.9944	A	*449.2354*	225.1214	432.2089	216.6081	431.2249	216.1161	5
21	2020.0242	1010.5157	2002.0136	1001.5105	S	*378.1983*	189.6028	361.1718	181.0895	360.1878	180.5975	4
22	2077.0457	1039.0265	2059.0351	1030.0212	G	291.1663	146.0868	274.1397	137.5735	273.1557	137.0815	3
23	2164.0777	1082.5425	2146.0671	1073.5372	S	234.1448	117.5761	217.1183	109.0628	216.1343	108.5708	2
24					K	147.1128	74.0600	130.0863	65.5468			1

B

Precursor Ion = 1163.5893 (2+)
Mascot Score = 131
Measured Mass = 2325.1641
Calculated Mass = 2325.1709
ppm = - 2.92

CLPP3-STREPII



m/z

ADVSTVCLGLAAS$_{164}$MGAFLLASGSK

#	b	b++	b0	b0++	Seq.	y	y++	y*	y*++	y0	y0++	#
1	72.0444	36.5258			A							24
2	187.0713	94.0393	169.0608	85.0340	D	2255.1410	1128.0741	2238.1145	1119.5609	2237.1305	1119.0689	23
3	286.1397	143.5735	268.1292	134.5682	V	2140.1141	1070.5607	2123.0875	1062.0474	2122.1035	1061.5554	22
4	373.1718	187.0895	355.1612	178.0842	S	2041.0457	1021.0265	2024.0191	1012.5132	2023.0351	1012.0212	21
5	474.2195	237.6134	456.2089	228.6081	T	*1954.0136*	977.5105	1936.9871	968.9972	1936.0031	968.5052	20
6	573.2879	287.1476	555.2773	278.1423	V	1852.9660	926.9866	1835.9394	918.4733	1834.9554	917.9813	19
7	733.3185	367.1629	715.3080	358.1576	C	*1753.8975*	877.4524	1736.8710	868.9391	1735.8870	868.4471	18
8	846.4026	423.7049	828.3920	414.6996	L	*1593.8669*	797.4371	1576.8403	788.9238	1575.8563	788.4318	17
9	903.4240	452.2157	885.4135	443.2104	G	*1480.7828*	740.8951	1463.7563	732.3818	1462.7723	731.8898	16
10	1016.5081	508.7577	998.4975	499.7524	L	1423.7614	712.3843	1406.7348	703.8710	1405.7508	703.3790	15
11	1087.5452	544.2762	1069.5347	535.2710	A	*1310.6773*	655.8423	1293.6508	647.3290	1292.6667	646.8370	14
12	1158.5823	579.7948	1140.5718	570.7895	A	*1239.6402*	620.3237	1222.6136	611.8105	1221.6296	611.3184	13
13	1245.6144	623.3108	1227.6038	614.3055	S	1168.6031	584.8052	1151.5765	576.2919	1150.5925	575.7999	12
14	1376.6549	688.8311	1358.6443	679.8258	M	*1081.5710*	541.2892	1064.5445	532.7759	1063.5605	532.2839	11
15	1433.6763	717.3418	1415.6657	708.3365	G	*950.5306*	475.7689	933.5040	467.2556	932.5200	466.7636	10
16	1504.7134	752.8604	1486.7029	743.8551	A	*893.5091*	447.2582	876.4825	438.7449	875.4985	438.2529	9
17	1651.7818	826.3946	1633.7713	817.3893	F	*822.4720*	411.7396	805.4454	403.2264	804.4614	402.7343	8
18	1764.8659	882.9366	1746.8553	873.9313	L	*675.4036*	338.2054	658.3770	329.6921	657.3930	329.2001	7
19	1877.9500	939.4786	1859.9394	930.4733	L	*562.3195*	281.6634	545.2930	273.1501	544.3089	272.6581	6
20	1948.9871	974.9972	1930.9765	965.9919	A	*449.2354*	225.1214	432.2089	216.6081	431.2249	216.1161	5
21	2036.0191	1018.5132	2018.0085	1009.5079	S	*378.1983*	189.6028	361.1718	181.0895	360.1878	180.5975	4
22	2093.0406	1047.0239	2075.0300	1038.0186	G	291.1663	146.0868	274.1397	137.5735	273.1557	137.0815	3
23	2180.0726	1090.5399	2162.0620	1081.5347	S	234.1448	117.5761	217.1183	109.0628	216.1343	108.5708	2
24					K	147.1128	74.0600	130.0863	65.5468			1

3.3.2 The catalytic activity of CLPP5 is required for function; S-to-A change in CLPP5 prevents complementation of clpp5-1

Similar as for CLPP3, heterozygous *clpp5-1* was transformed with genomic *CLPP5* with a C-terminal STREPII tag (*CLPP5-STREPII*), or the same construct but with a serine to alanine mutation in the catalytic site (*CLPP5S193A-STREPII*) that rendered CLPP5 catalytically inactive. Expression of *CLPP5-STREPII* could fully complement the *clpp5-1* null mutant (Figure 2A - middle panel) but *CLPP5S193A-STREPII* could not. We also identified several lines expressing *CLPP5S193A-STREPII* in wt background or in the heterozygous *clpp5-1* background (Supplemental Figure 2A,D); these lines did not show any visible phenotype (Figure 2A). The S193A mutation was confirmed by sequencing of RT-PCR products, as well at the protein level by MS/MS-based identification of peptides covering this catalytic residue (Figure 2B,C; Supplemental Figure 4). Furthermore, similar as we determined for CLPP3, STREPII-tagged CLPP5 and CLPP5S193A accumulated at comparable levels (Figure 2D) and normally assembled into CLPRT complexes (Figure 2E). The fact that *clpp5-1(Aa)/CLPP5S193A-STREPII* does not have a visible phenotype indicates that accumulation of the catalytically inactive CLPP5 in the presence of endogenous CLPP5 does not reduce the CLP protease capacity below the minimum threshold level required for chloroplast biogenesis, proteostasis, and function.

Despite extensive screening efforts, no homozygous *clpp5-1* lines expressing *CLPP5S193A-STREPII* were identified, in contrast to the many homozygous *clpp5-1* lines expressing *CLPP5-STREPII.* This suggests that complete loss of catalytic activity of CLPP5 results in embryo lethality. To further test this hypothesis, we

analyzed the segregation pattern of the phenotypes of developing seeds in the siliques of wt, wt expressing *CLPP5S193A-STREPII,* heterozygous *clpp5-1*, and heterozygous *clpp5-1* expressing *CLPP5S193A-STREPII*. The first two lines produced only green seeds in developing siliques, while heterozygous *clpp5-1* made both green and white seeds in a 3:1 ratio (Figure 2F), similar as previously observed (*20*). Importantly, ~3:1 segregating ratios were also found in the progeny of heterozygous *clpp5-1* carrying *CLPP5S193A-STREPII* (Figure 2F,G). This confirms that the serine to alanine change of CLPP5 prevented complementation of *clpp5-1* null mutants, and thus that the complete loss of catalytic activity of CLPP5 which reduces the number of catalytic triads per complex from 10 to 7, results in embryo lethality.

3.3.3 Affinity-purification of CLPP3 and CLPP5 STREPII-tagged complexes

To identify potential substrates using the CLP core trapping technique as explained in the INTRODUCTION, we carried out replicate affinity experiments for homozygous *clpp3-1* complemented with *CLPP3S164A-STREPII* and using *clpp3-1* complemented by *CLPP3-STREPII* as a control. In the case of CLPP5, we used heterozygous *clpp5-1* expressing *CLPP5S193A-STREPII*, with homozygous *clpp5-1* expressing *CLPP5-STREPII* as the control. The STREPII-tagged complexes were purified on streptactin columns. The affinity eluates were each run out on SDS-PAGE gels, and each gel lane was cut into gel slices, followed by digestion with trypsin and MS/MS analysis for protein identification.

First, we directly (*i.e.* eluates were run on the same SDS-PAGE gel and processed in parallel) compared a negative control affinity purification using soluble protein extracts of wt plants (*i.e*. these lack STREPII-tagged proteins) with CLPP3-

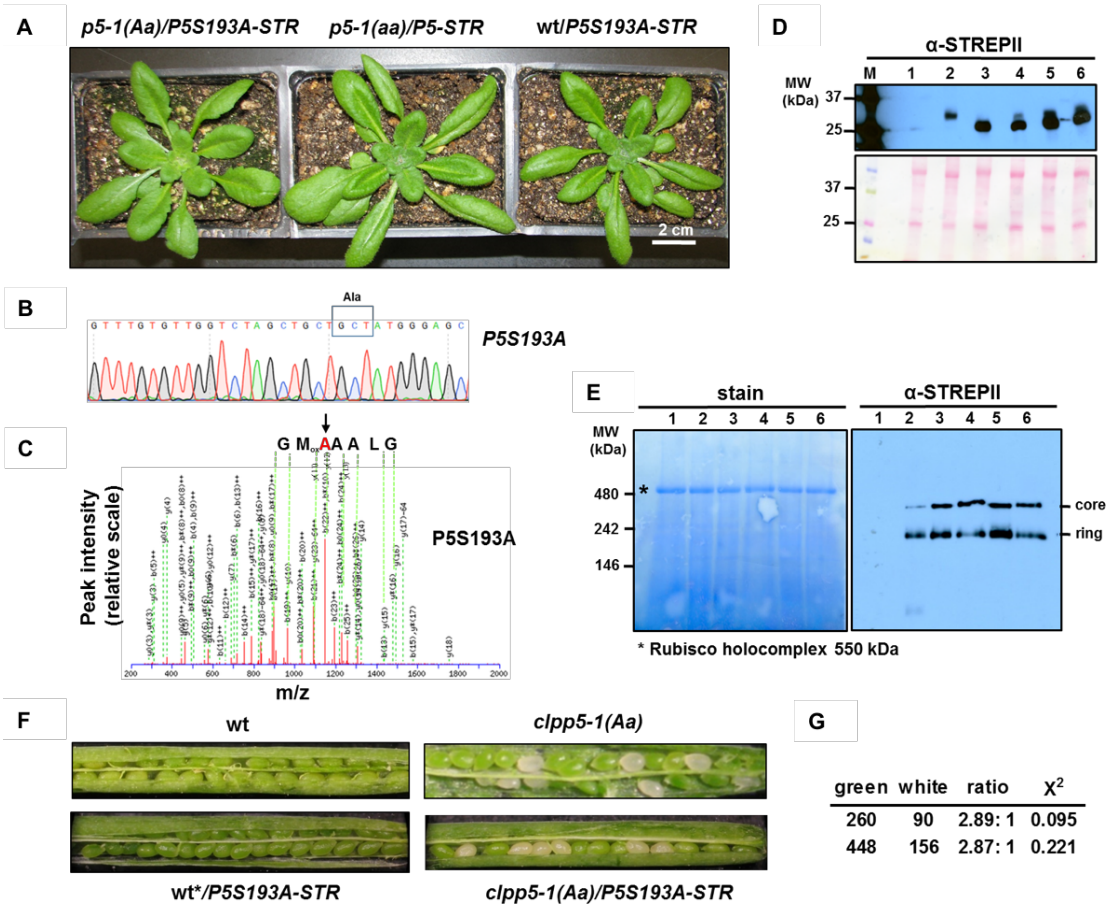


Figure 2. Characterization of wt and *clpp5-1* lines expressing *CLPP5S193A-STREPII* or *CLPP5-STREPll* transgenes.

A, Phenotypic analysis of wt and heterozygous (Aa) *clpp5-1* expressing *CLPP5S193A-STREPll* or homozygous *clpp5-1* expressing *CLPP5-STREPII* grown on soil for 24 days under 10/14 hr light/dark cycle at 120-150 µmol photons.m^{-2}.s^{-1}. **B,** Confirmation of the point mutation in the catalytic sites of *CLPP5S193A-STREPII* by DNA sequencing of RT-PCR products (primers #14 and #9) amplified from transgenic plants. The codon responsible for the Ser to Ala point mutation is boxed and is AGT in wt but GCT in the mutant. **C,** Confirmation of the point mutation in the catalytic sites of CLPP5S193A-STREPII by MS/MS of the tryptic peptide (HIRPDVSTVCVGLAAA$_{193}$MGAFLLSAGTK) generated by tryptic digestion of affinity purified CLP complexes. The MS/MS spectrum is from a triply charged

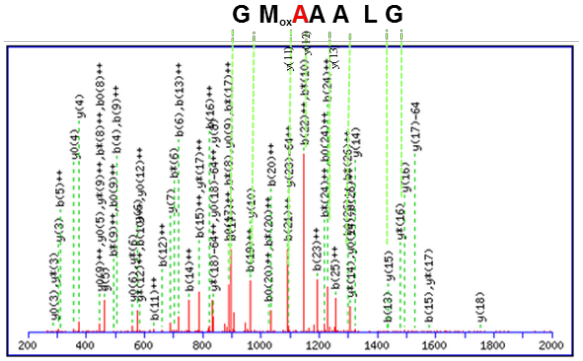
precursor ion with m/z of 914.8238 (3^+) with MASCOT ion score of 83 (-5 ppm error) and supports the residue A164. The partial peptide sequence listed above the spectrum ($GMA_{193}AALGL$) shown is based on y-ions explaining this reads in reverse order. A list of b- and y-ions is provided in Supplemental Figure 4A. An example of an MS/MS spectrum of the analogous CLPP5 wild-type peptide is provided in Supplemental Figure 4B. **D,** SDS-PAGE and immunoblotting of total soluble protein from wt (lane 1), *clpp3-1* with *CLPP3-STREPIl* (lane 2), *clpp5-1* with *CLPP5-STREPII* (lane 3), and *clpp5-1* heterozygous mutant with *CLPP5S193A-STREPII* (lanes 4-6). The Ponceau red stain of the blot is shown. Anti-STREPII serum was used. **E,** Native gels and immunoblotting of total leaf protein from wt (lane 1), *clpp3-1* with *CLPP3-STREPII* (lane 2), *clpp5-1* with *CLPP5-STREPII* (lanes 3 and 5), the *clpp5-1* heterozygous mutant with *CLPP5S193A-STREPII* (lanes 4 and 6). The CLPPRT core and ring are indicated. Anti-STREPII serum was used. * indicates the Rubisco complex. **F,** Developing siliques of wt, heterozygous *clpp5-1 (Aa),* heterozygous *clpp5-1 (Aa)* with *CLPP5S193A-STREPll,* and wt* with *CLPP5S193A-STREPII* obtained from a prior segregating progeny of heterozygous *clpp5-1* with *CLPP5S193A-STREPll.* The segregating white seeds are indicative of impaired chloroplast development due to the homozygous *clpp5-1* background. **G,** Segregation analysis of green and white seeds in developing siliques of two heterozygous *clpp5-1* mutants expressing *CLPP5S193A-STREPII.*

Supplemental Figure 4. MS/MS-based verification of the S193A mutation in CLPP5S193A-STREPII and comparison to CLPP5-WT-STREPII

A, Confirmation of the point mutation in the catalytic sites of CLPP5S193A-STREPII by MS/MS of the tryptic peptide (HIRPDVSTVCVGLAAAMGAFLLSAGTK) generated by tryptic digestion of affinity purified CLP complexes. The MS/MS spectrum is from a triply charged precursor ion with m/z of 914.8238 (3^+) with MASCOT ion score of 83 (-5 ppm error) and supports the residue A193. A list of b- and y-ions is listed. **B,** An example of an MS/MS spectrum of the CLPP5 wild-type peptide (HIRPDVSTVCVGLAASMGAFLLSAGTK) covering the region around S193**.** The MS/MS spectrum is from a triply charged precursor ion with m/z of 920.150193 (3^+) with MASCOT ion score of 126 (0.68 ppm error). The partial peptide sequence listed above the spectrum ($GMS_{193}AALG$) shown is based on y-ions explaining this reads in reverse order. A list of b- and y-ions is listed.

CLPP5-S193A-STREPII

A



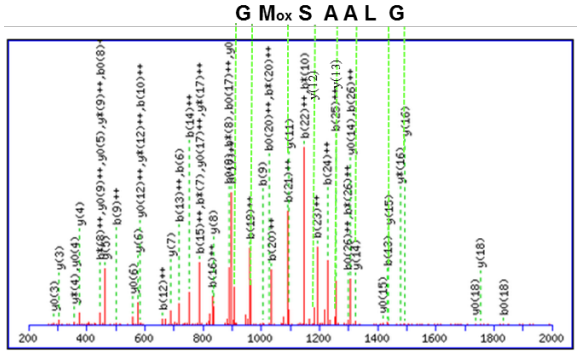
Precursor Ion = 914.8238 (3+)
Mascot Score = 83
Measured Mass = 2741.43575
Calculated Mass = 2741.4495
ppm = -5.0

HIRPDVSTVCVGLAAA$_{193}$MGAFLLSAGTK

#	b	b^{++}	b*	b*$^{++}$	b^{0}	b^{0++}	Seq.	y	y^{++}	y*	y*$^{++}$	y^{0}	y^{0++}	#
1	138.0662	69.5367					H							27
2	251.1503	126.0788					I	2621.3790	1311.1931	2604.3524	1302.6798	2603.3684	1302.1878	26
3	407.2514	204.1293	390.2248	195.6160			R	2508.2949	1254.6511	2491.2683	1246.1378	2490.2843	1245.6458	25
4	504.3041	252.6557	487.2776	244.1424			P	2352.1938	1176.6005	2335.1672	1168.0873	2334.1832	1167.5952	24
5	619.3311	*310.1692*	602.3045	301.6559	601.3205	301.1639	D	2255.1410	1128.0741	2238.1145	1119.5609	2237.1305	1119.0689	23
6	718.3995	359.7034	*701.3729*	351.1901	700.3889	350.6981	V	2140.1141	1070.5607	2123.0875	1062.0474	2122.1035	1061.5554	22
7	805.4315	403.2194	788.4050	394.7061	787.4209	394.2141	S	2041.0457	1021.0265	2024.0191	1012.5132	2023.0351	1012.0212	21
8	906.4792	453.7432	*889.4526*	*445.2300*	*888.4686*	*444.7380*	T	1954.0136	977.5105	1936.9871	968.9972	1936.0031	968.5052	20
9	1005.5476	*503.2774*	988.5211	494.7642	987.5370	494.2722	V	1852.9660	926.9866	1835.9394	918.4733	1834.9554	917.9813	19
10	1165.5783	*583.2928*	*1148.5517*	574.7795	1147.5677	574.2875	C	*1753.8975*	877.4524	1736.8710	868.9391	1735.8870	868.4471	18
11	1264.6467	*632.8270*	1247.6201	624.3137	1246.6361	623.8217	V	1593.8669	797.4371	*1576.8403*	*788.9238*	1575.8563	788.4318	17
12	1321.6681	*661.3377*	1304.6416	652.8244	1303.6576	652.3324	G	*1494.7985*	747.9029	*1477.7719*	739.3896	1476.7879	738.8976	16
13	*1434.7522*	*717.8797*	1417.7256	709.3665	1416.7416	708.8745	L	*1437.7770*	719.3921	1420.7505	710.8789	1419.7665	710.3869	15
14	1505.7893	*753.3983*	1488.7628	744.8850	1487.7787	744.3930	A	*1324.6930*	662.8501	*1307.6664*	654.3368	*1306.6824*	653.8448	14
15	*1576.8264*	*788.9168*	1559.7999	780.4036	1558.8159	779.9116	A	1253.6558	627.3316	1236.6293	618.8183	1235.6453	618.3263	13
16	1647.8635	*824.4354*	1630.8370	815.9221	1629.8530	815.4301	A	1182.6187	591.8130	1165.5922	*583.2997*	1164.6082	*582.8077*	12
17	1794.8989	*897.9531*	1777.8724	*889.4398*	1776.8884	*888.9478*	M	1111.5816	556.2944	1094.5551	547.7812	1093.5710	547.2892	11
18	1851.9204	926.4638	1834.8939	917.9506	1833.9098	917.4586	G	*964.5462*	482.7767	947.5197	474.2635	946.5356	473.7715	10
19	1922.9575	*961.9824*	1905.9310	953.4691	1904.9470	952.9771	A	907.5247	454.2660	890.4982	*445.7527*	*889.5142*	*445.2607*	9
20	2070.0259	*1035.5166*	2052.9994	*1027.0033*	2052.0154	*1026.5113*	F	*836.4876*	418.7475	819.4611	410.2342	818.4771	409.7422	8
21	2183.1100	*1092.0586*	2166.0834	1083.5454	2165.0994	1083.0534	L	*689.4192*	345.2132	672.3927	336.7000	671.4087	336.2080	7
22	2296.1941	*1148.6007*	2279.1675	1140.0874	2278.1835	1139.5954	L	*576.3352*	288.6712	*559.3086*	280.1579	*558.3246*	279.6659	6
23	2383.2261	*1192.1167*	2366.1995	1183.6034	2365.2155	1183.1114	S	*463.2511*	232.1292	446.2245	223.6159	*445.2405*	223.1239	5
24	2454.2632	*1227.6352*	2437.2367	*1219.1220*	2436.2526	*1218.6300*	A	*376.2191*	188.6132	359.1925	180.0999	*358.2085*	179.6079	4
25	2511.2847	*1256.1460*	2494.2581	1247.6327	2493.2741	1247.1407	G	*305.1819*	153.0946	*288.1554*	144.5813	*287.1714*	144.0893	3
26	2612.3323	*1306.6698*	2595.3058	*1298.1565*	2594.3218	*1297.6645*	T	248.1605	124.5839	231.1339	116.0706	230.1499	115.5786	2
27							K	147.1128	74.0600	130.0863	65.5468			1

CLPP5-STREPII

B



Precursor Ion = 920.150193 (3+)
Mascot Score = 126
Measured Mass = 2757.428751
Calculated Mass = 2757.4306
ppm = 0.68

HIRPDVSTVCVGLAAS$_{193}$MGAFLLSAGTK

#	b	b^{++}	b*	b*$^{++}$	b^{0}	b^{0++}	Seq.	y	y^{++}	y*	y*$^{++}$	y^{0}	y^{0++}	#
1	138.0662	69.5367					H							27
2	251.1503	126.0788					I	2621.3790	1311.1931	2604.3524	1302.6798	2603.3684	1302.1878	26
3	407.2514	204.1293	390.2248	195.6160			R	2508.2949	1254.6511	2491.2683	1246.1378	2490.2843	1245.6458	25
4	504.3041	252.6557	487.2776	244.1424			P	2352.1938	1176.6005	2335.1672	1168.0873	2334.1832	1167.5952	24
5	619.3311	310.1692	602.3045	301.6559	601.3205	301.1639	D	2255.1410	1128.0741	2238.1145	1119.5609	2237.1305	1119.0689	23
6	*718.3995*	359.7034	701.3729	351.1901	700.3889	350.6981	V	2140.1141	1070.5607	2123.0875	1062.0474	2122.1035	1061.5554	22
7	805.4315	403.2194	*788.4050*	394.7061	787.4209	394.2141	S	2041.0457	1021.0265	2024.0191	1012.5132	2023.0351	1012.0212	21
8	906.4792	453.7432	*889.4526*	*445.2300*	*888.4686*	*444.7380*	T	1954.0136	977.5105	1936.9871	968.9972	1936.0031	968.5052	20
9	*1005.5476*	*503.2774*	988.5211	494.7642	987.5370	494.2722	V	1852.9660	926.9866	1835.9394	918.4733	1834.9554	917.9813	19
10	1165.5783	*583.2928*	*1148.5517*	574.7795	1147.5677	574.2875	C	*1753.8975*	877.4524	1736.8710	868.9391	*1735.8870*	868.4471	18
11	1264.6467	632.8270	1247.6201	624.3137	1246.6361	623.8217	V	1593.8669	797.4371	1576.8403	*788.9238*	1575.8563	*788.4318*	17
12	1321.6681	*661.3377*	1304.6416	652.8244	1303.6576	652.3324	G	*1494.7985*	747.9029	*1477.7719*	739.3896	1476.7879	738.8976	16
13	*1434.7522*	*717.8797*	1417.7256	709.3665	1416.7416	708.8745	L	*1437.7770*	719.3921	1420.7505	710.8789	*1419.7664*	710.3869	15
14	1505.7893	*753.3983*	1488.7628	744.8850	1487.7787	744.3930	A	*1324.6930*	662.8501	1307.6664	654.3368	*1306.6824*	653.8448	14
15	1576.8264	*788.9168*	1559.7999	780.4036	1558.8159	779.9116	A	1253.6558	627.3316	1236.6293	618.8183	1235.6453	618.3263	13
16	1663.8585	*832.4329*	1646.8319	823.9196	1645.8479	823.4276	S	1182.6187	591.8130	1165.5922	*583.2997*	1164.6082	*582.8077*	12
17	1794.8989	*897.9531*	1777.8724	*889.4398*	1776.8884	*888.9478*	M	*1095.5867*	548.2970	1078.5601	539.7837	1077.5761	539.2917	11
18	1851.9204	926.4638	1834.8939	917.9506	*1833.9098*	917.4586	G	964.5462	482.7767	947.5197	474.2635	946.5356	473.7715	10
19	1922.9575	*961.9824*	1905.9310	953.4691	1904.9469	952.9771	A	907.5247	454.2660	890.4982	*445.7527*	*889.5142*	*445.2607*	9
20	2070.0259	*1035.5166*	2052.9994	*1027.0033*	2052.0154	*1026.5113*	F	*836.4876*	418.7475	819.4611	410.2342	818.4771	409.7422	8
21	2183.1100	*1092.0586*	2166.0834	1083.5454	2165.0994	1083.0534	L	*689.4192*	345.2132	672.3927	336.7000	671.4087	336.2080	7
22	2296.1941	*1148.6007*	2279.1675	1140.0874	2278.1835	1139.5954	L	*576.3352*	288.6712	559.3086	280.1579	*558.3246*	279.6659	6
23	2383.2261	*1192.1167*	2366.1995	1183.6034	2365.2155	1183.1114	S	*463.2511*	232.1292	446.2245	223.6159	*445.2405*	223.1239	5
24	2454.2632	*1227.6352*	2437.2366	1219.1220	2436.2526	1218.6300	A	*376.2191*	188.6132	*359.1925*	180.0999	*358.2085*	179.6079	4
25	2511.2847	*1256.1460*	2494.2581	1247.6327	2493.2741	1247.1407	G	*305.1819*	153.0946	288.1554	144.5813	287.1714	144.0893	3
26	2612.3323	*1306.6698*	2595.3058	*1298.1565*	2594.3218	*1297.6645*	T	248.1605	124.5839	231.1339	116.0706	230.1499	115.5786	2
27							K	147.1128	74.0600	130.0863	65.5468			1

STREPII and CLPP3S164A-STREPII purifications (Supplemental Table 2). Biotin-containing proteins bind to streptactin columns and Arabidopsis contains three major endogenous biotin-binding complexes, namely cytosolic (ACC1) and plastidic acetyl-CoA carboxylase complexes (ACCase), mitochondrial 3-methylcrotonyl-CoA carboxylase complexes (MCC) (*40*). Since the concentration of these complexes should be similar in wt and CLPP3-STREPII lines, they serve as internal controls. We detected high numbers of MS/MS spectra for these endogenous biotin binding proteins (ACC1, MCCA, MCCB, and BIOTIN CARBOXYLASE (BC) - part of the ACCase) (Figure 3A). Indeed, direct comparison between affinity eluates of wt and the two CLPP3-STREPII lines showed very similar levels of these endogenous biotin-containing complexes (average ratio 1.34) (Figure 3A). This provides an internal calibrant for evaluating proteins enriched in the affinity eluates. In contrast to these endogenous biotin binding complexes, CLP subunits were on average 25-fold higher in the CLPP3-STREPII lines than wt plants (Figure 3A), showing that the affinity purification of CLP complexes worked well.

A total of 16 successful independent affinity experiments for CLPP3 (four for each CLPP3-STREPII and CLPPP3S164A-STREPII) and CLPP5 (four for each CLPP5-STREPII and CLPP5S193A-STREPII) were carried out. We note that for all CLPP5 affinity purifications, we reduced the level of co-purifying endogenous biotin binding proteins by preincubation of total soluble leaf extract with avidin; yet, there was still sufficient endogenous biotin binders for internal calibration (Supplemental Figure 5). The identified proteins, their annotation and number of matched MS/MS spectra (adjSPC – see METHODS for explanation) for each experiment are assembled

in Supplemental Table 2. All chloroplast CLPPRT core subunits including CLPP (P1, P3-6), CLPR (R1-4), as well as CLPT1,2 subunits were identified. These subunits of the CLP complex represented on average 19% of all adjSPC matched to plastid proteins in the case of CLPP3-STREPII and CLPP3S164A-STREPII, and on average 37% in the case of CLPP5-STREPII and CLPP5S193A-STREPII (Figure 3B). This shows that the serine to alanine mutation did not affect the yield of the CLPPR core complexes. We also identified low levels of CLPC1,2 chaperones in 14 out of 16 experiments without significant differences between the four STREPII tagged lines (Supplemental Table 2). We did not identify the adaptor CLPS1 and only once adaptor CLPF (with 3 MS/MS spectra in a CLPP5S193A experiment). The lack of consistent observation of these adaptors is consistent with their substrate delivery role and transient interaction to CLPC chaperones (*25, 41*). Together, these affinity experiments indicate that i) the loss of catalytic triads in CLPP3 or CLPP5 did not affect the yield of CLP core affinity purification, complex composition or interaction between the CLPC and the CLP core, and that ii) the CLPP5-STREPII affinity purification was ~2-fold more efficient than the CLPP3-STREPII purification (Figure 3B), likely because there were on average more CLPP5-STREPII copies per CLP core complex as illustrated in Figure 3C.

3.3.4 Trapping substrates in affinity-purified CLPP3S164A-STREPII and CLPP5S193A-STREPII complexes

We then evaluated the affinity experiments for other (non-CLP) chloroplast proteins enriched in these CLPP3- and CLPP5 STREPII-tagged complexes; these could be interactors to the CLP core or CLP substrates located within the chamber of

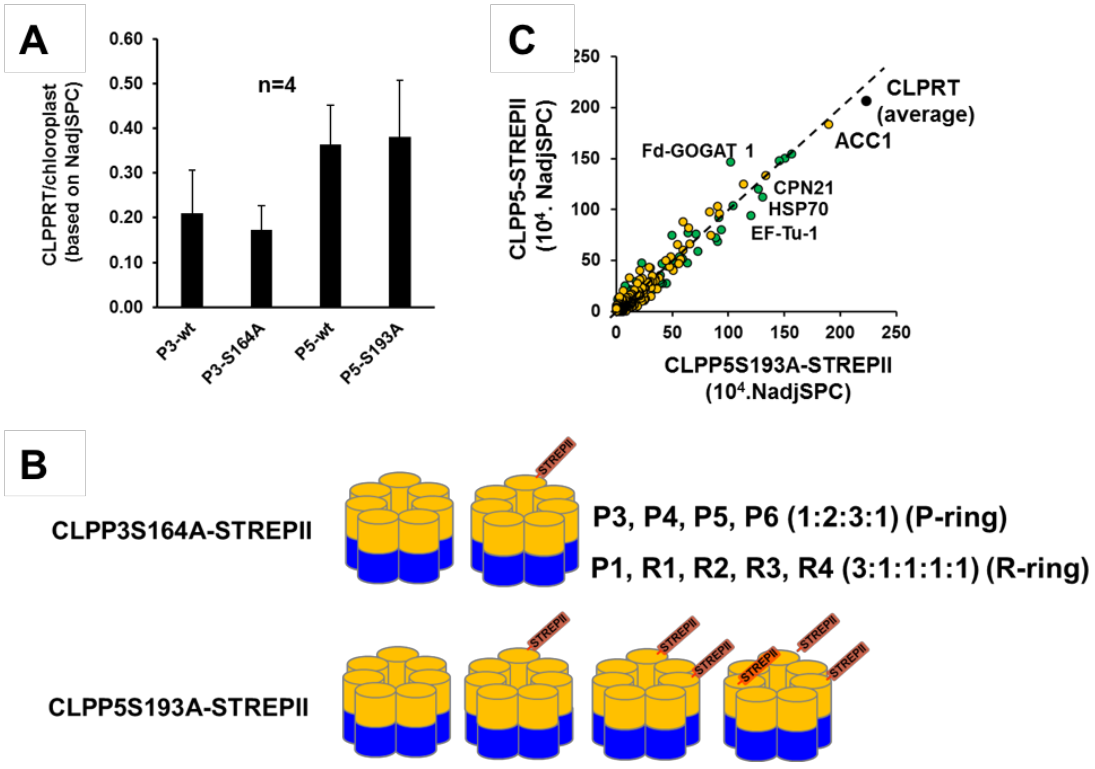
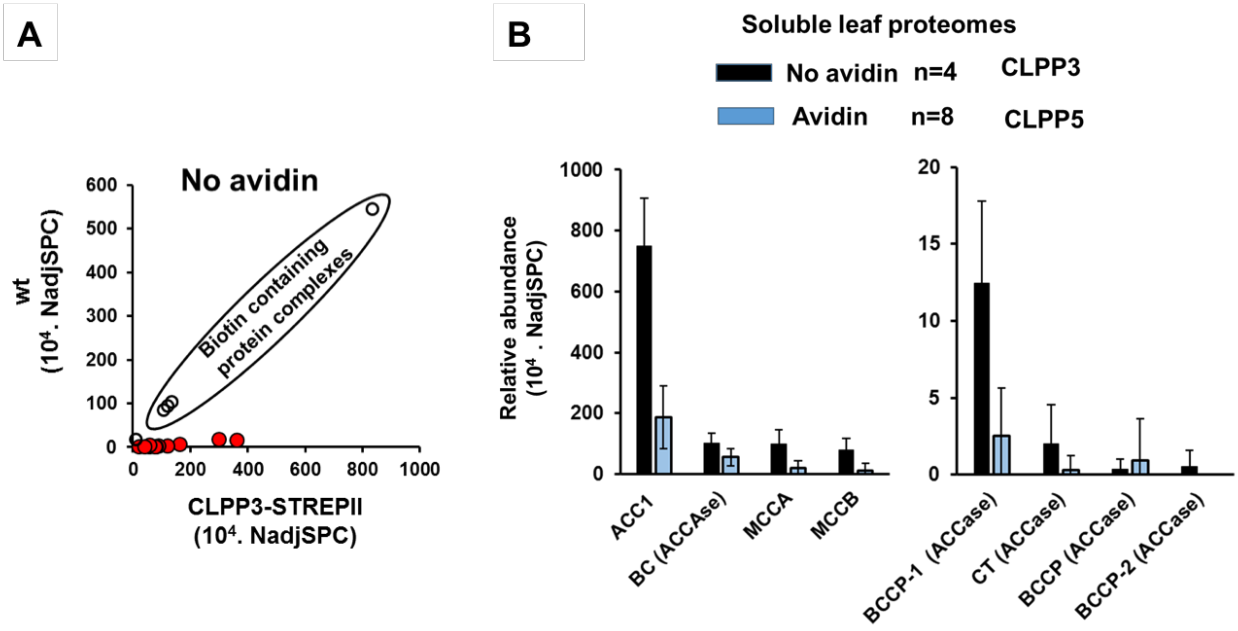


Figure 3. Affinity purification of STREPII-tagged CLP core complexes and MS/MS analysis.

A, Relative abundance of endogenous Arabidopsis biotin-binding ACC1, MCCA, BIOTIN CARBOXYLASE (BC) and MCCB, and subunits of the CLPPR complex in streptactin affinity eluates from wt (col-o) plants and CLPP3-STREPII transgenic lines. CLP subunits are marked in red-filled circles. In these experiments, no avidin was added prior to affinity purifications; addition of avidin does reduce the binding of endogenous biotin protein proteins (see Supplemental Figure 5). **B,** The average yield of CLPPRT subunits for affinity-purified complexes containing CLPP3-STREPII, CLPP3S164A-STREPII, CLPP5-STREPII and CLPP5S193A-STREPII (n=4 for each transgene). Yields were normalized to the total amount of detected plastid proteins, using NadjSPC as the basis for quantification. **C,** Schematic representation of the wt and tagged CLPPR core complexes with P-rings and R-ring. Both ClpP3 and CLPP5 are part of the P-ring, with respectively one or three copies per ring in native complexes (*15*). In the case of the STREPII-tagged CLPP3 and CLPP3S164A lines,

each complex has one STREPII-tagged CLPP3 protein and no endogenous CLPP3. However in case of *CLPP5S193A-STREPII* lines, plants have both endogenous CLPP5 and CLPP5S193A-STREPII; affinity purified CLPP5 complexes from these lines have one, two or three CLPP5S193A-STREPII subunits per P-ring. We note that the transgenes are genomic constructs with their endogenous promoters. The positions of the STREPII tagged proteins within the P-ring are unknown. **D,** Cross-correlation of proteins of average abundance (based on NadjSPC) identified in affinity eluates of *CLPP5-STREPII* and *CLPP5S193A-STREPII* lines. These proteins were observed two or more times across the eight preparations (4 of each STREPII line). Proteins localized in the chloroplasts are indicated in green and proteins located outside the chloroplast or without annotated subcellular locations are indicated in yellow; subcellular location was based on annotation of manually curated experimental evidence collected from information in the public domain and in-house data (from PPDB). The data point labeled 'CLPPRT' is the average of all 11 CLPP, CLPR and CLPT subunits. ACC1 (ACETYL-COA-CARBOXYLASE 1; AT1G36160) has biotin as the cofactor and has therefore high affinity to the streptactin column. HSP70-1 (AT4G24280); HSP70-2 (AT5G49910); RAC – RUBISCO ACTIVASE (AT2G39730); CPN21 (AT5G20720); EF-TU - ELONGATION FACTOR TU (AT4G20360); THI1- THIAMIN BIOSYNTHESIS 1 (AT5G54770).



Supplemental Figure 5. Effect of preincubation of soluble leaf proteomes with avidin to reduce the binding and enrichment on streptactin columns for CLPP3(S164A)-STREPII and CLPP5(S193A)-STREPII lines.

The black bars indicate the abundances of the co-purified biotin-containing proteins from CLPP3(S164A)-STREPII soluble leaf proteomes without avidin pretreatment. n=4 includes two CLPP3-STREPII and two CLPP3S164A-STREPII lines. The blue bars show those from the avidin-pretreated CLPP5(S193A)-STREPII soluble leaf proteomes. n=8 includes four CLPP5-STREPII and four CLPP5S193A-STREPII lines. Standard deviations are indicated.

the protease core. The theoretical basis for trying to identify substrates in CLP cores is based on the hypothesis that chloroplast CLP complexes that lack one or more of the 10 catalytic sites (three in the R ring and seven in the P ring) have slower catalytic rates and therefore the average time to degrade substrates is extended; consequently such substrates reside for a longer time in the CLP core complex. MS/MS of such affinity enriched complexes could allow recognition of such substrates as having increased abundance when compared to fully active affinity-enriched complexes. In the case of CLPP5, there are three endogenous copies per core (*15*). Since we expressed the catalytically inactive CLPP5S193ASTREPII in the heterozygous *clpp5-1* mutant, there are both endogenous CLPP5 subunits and inactive tagged CLPP5 copies. We note that the CLPP3 and CLPP5 transgenes are generated from genomic sequences including the endogenous promoters, as to ensure more natural levels of expression. The affinity purified STREPII-tagged CLP cores must have either one, or possibly two or three inactive CLPP5 subunits (Figure 3C); in particular those complexes with 2 or 3 inactive CLPP5 subunits could have reduced CLP catalytic rates. However, since it is quite likely that the CLP protease is constantly involved in monitoring and degrading chloroplast proteins, it is possible that also fully active CLP core complexes contain substrates.

We compared the eluates of the wt and mutant forms of CLPP3- and CLPP5-STREPII-tagged complexes based on the normalized number of matched MS/MS spectra (NadjSPC) (Supplemental Table 2). We tested two types of thresholds (*34-36, 42, 43*) to identify possible candidate substrates: i) only observed in eluates of mutant complexes and not in CLPP3/5-STREPII controls, or ii) at least 2-fold enriched in

eluates of the mutant complexes compared to CLPP3/5-STREPII controls. To reduce stochastic noise, we applied also either a minimum frequency of observation (2 or 3) across replicates or an abundance threshold of 1.10^{-3} based on NadjSPC (this corresponds to ~0.1% of protein mass).

In the case of CLPP3, we isolated chloroplasts and extracted soluble stromal proteomes and carried out two independent affinity purifications for each CLPP3-STREPII and CLPP3S164A-STREPII (trap) (Supplemental Table 2). 17 proteins were only found in the mutant and 12 proteins were >2x fold enriched in the mutant eluates; however, because the number of matched MS/MS spectra were low or variable across replicates, we did not consider them strong candidate substrates. The exception was GLUTAMINE SYNTHASE 2 (GS2; AT5G35630), because it was observed in both CLPP3S164A-STREPII eluates with either three or six MS/MS spectra, but not in CLPP3-STREPII. For comparison we note that the average ratio between CLP core subunits between CLPP3S164A-STREPII and CLPP3-STREPII was 0.7 and the ratio for stromal biotin carboxylase (BC) was 0.77. In a second set of CLPP3 experiments, we used soluble total leaf extracts and a total of four affinity purifications (two mutants and two controls). The average ratio between CLP core subunits between CLPP3S164A-STREPII and CLPP3-STREPII was 1.0 and the ratio for endogenous biotin binder and internal control ACCase was 1.2. We applied these same thresholds (only in the mutant P3 or >2x enriched) and identified nine proteins that were more than 2-fold enriched; however, following closer inspection we do not suggest them to be strong candidates, based on the same arguments as for the stromal replicates (Supplemental Table 2).

In the case of CLPP5, our dataset consists of four affinity experiments with each CLPP5S193A-STREPII and CLPP5-STREPII and using total soluble leaf extracts (Supplemental Table 2). Figure 3D shows a cross-correlation of the average relative protein abundances (based on NadjSPC) across the replicate affinity purified CLPP5S193A-STREPII and CLPP5-STREPII complexes. This shows that the plastid and non-plastid proteins co-purified with the fully active and partially active complexes have generally similar relative abundances. Only one plastid protein was enriched, namely FRUCTOSE-BISPHOSPHATASE (FBPA; HCEF1; AT3G54050) (*44*) with an average ratio of 2.2 (standard deviation 1.1). Interesting, the abundant stromal proteins THI1 and EF-TU were previously suggested to be CLP substrates (*4, 45*) and were observed with high abundance in the eluates of both CLPP5-STREPII and CLPP5S193A-STREPII (Figure 3D). Furthermore stromal CPN21, abundant in eluates of both CLPP5 constructs (mutant/control ratio 1.2 ±1.2) (Figure 3D), was previously identified as a strong interactor in CLP affinity purification using CLPT1-STREPII and CLPT2-STREPII Arabidopsis lines (*17*).

3.3.5 The bottleneck for degradation is likely substrate recognition and unfolding by CLP adaptors and chaperones, upstream of the CLP core

We showed that CLPP3 (one copy per CLP core) makes an essential structural contribution, but that its catalytic site is dispensable for plant growth and development, whereas the catalytic activity of CLPP5 (3 copies per CLP core) is essential. Based on the extensive affinity experiments described in this study, we conclude that the CLP core catalytic activity in the CLPP3S164A-STREPII and CLPP5 S193A-STREPII complexes with reduced number of catalytic triads did not

result in significant accumulation of substrates as compared to fully active CLP complexes. This suggests that the bottleneck for degradation is likely substrate recognition and unfolding by CLP adaptors and chaperones, upstream of the CLP core. *In vivo* substrate trapping through partial inactivation of unfolding activity of the CLPC/D chaperones could provide an alternative strategy for identification of candidate substrates. This has been a successful approach for the CLPC homolog in the gram-positive bacterium *Staphylococcus aureus* where the CLP system is not essential for viability (*46*).

3.4 METHODS

3.4.1 Plant materials, cloning, transformation, Genotyping, RT-PCR

The T-DNA insertion lines for *CLPP3* (At1g66670) and *CLPP5* (At1g02560) are SALK_065330 (*clpp3-1*) and SALK_007708 (*clpp5-1*), respectively, as described in (*20*). Heterozygous *clpp3-1* and *clpp5-1* were used for complementation. To generate transgene constructs, 3.5 kb of *CLPP3* and 3.7 kb of *CLPP5* genomic DNA including 1 kb upstream and 800 bp downstream from the coding region were cloned using gene-specific primers (primer set 1 – see Supplemental Table 1). A 24-nucleotide sequence, for the 8-amino acid strepll tag (WSHPQFEK), was introduced into the transgene before the stop codon by PCR amplification (primer set 2 – see Supplemental Table 1). To create the *CLPP3S164A-STREPll* and *CLPP5S193A-STREPII* constructs, the catalytic serine of CLPP3 and CLPP5 was substituted by alanine using PCR amplification using a primer set 3. All PCR products were sub-

Supplemental Table 2. Proteins identified across the 16 affinity experiments with CLPP3-STREPII, CLPP3S164A-STREPII, CLPP5-STREPII or CLPP5S193A-STREPII.

Accession (experiment # in PPDB: 2164 2165 2190 2191 2192 2193 2224 2225 2208 2209 2210 2211 2214 2215 2216 2217)	10000* NadjSPC exp. 2164 P3 control stroma	10000* NadjSPC exp. 2165 P3 S164A stroma	10000* NadjSPC exp. 2190 P3 control stroma	10000* NadjSPC exp. 2191 P3 S164A stroma	10000* NadjSPC exp. 2192 P3 S164A	10000* NadjSPC exp. 2193 P3 control	10000* NadjSPC exp. 2224 P3 control	10000* NadjSPC exp. 2225 P3 S164A	10000* NadjSPC Exp. 2208 P5 control	10000* NadjSPC exp. 2209 P5 S193A	10000* NadjSPC exp. 2210 P5 control	10000* NadjSPC exp. 2211 P5 S193A	10000* NadjSPC exp. 2214 P5 control	10000* NadjSPC exp. 2215 P5 S193A	10000* NadjSPC exp. 2216 P5 control	10000* NadjSPC exp. 2217 P5 S193A
ATCG00490.1	1025.4	2122.4	1375.5	894.0	1239.7	1307.3	621.4	862.3	1279.8	1144.4	887.2	1074.5	623.1	531.5	494.9	607.5
AT5G45390.1	398.8	348.9	268.4	241.1	296.4	237.9	312.8	302.1	460.9	650.8	627.7	307.6	610.1	661.4	524.3	742.2
AT1G02560.1	175.5	72.7	51.7	72.7	80.7	89.1	385.2	356.0	38.1	52.5	222.2	37.7	454.3	482.5	489.8	676.0
AT1G11750.1	239.3	130.8	80.5	170.6	129.4	137.6	159.6	175.2	384.7	380.4	648.7	389.6	242.9	233.3	334.2	438.5
AT5G54770.1	1754.6	1264.7	1187.6	1008.9	282.3	60.9	109.6	64.8	30.5	84.9	213.4	311.7	211.4	92.8	137.8	91.3
AT1G49970.1	193.7	130.8	174.4	115.0	56.5	19.7	99.0	67.6	129.5	490.5	188.3	94.3	254.1	429.5	159.4	306.0
AT2G39730.1	79.8	101.8	228.1	560.1	150.5	149.4	126.6	95.2	182.8	172.9	154.8	147.6	103.8	156.4	162.0	123.3
AT4G24280.1	153.8	203.5	252.3	383.5	78.3	89.2	82.4	101.5	101.3	154.1	155.7	147.2	224.9	165.4	136.7	157.1
AT1G09130.1	444.3	247.1	174.4	174.3	68.2	29.5	90.4	69.0	110.5	154.1	125.5	86.1	222.5	128.6	145.4	237.5
AT1G12410.1	159.5	145.4	40.3	85.3	16.5	11.8	105.3	74.5	175.2	194.9	146.5	127.1	168.7	182.9	187.5	230.7
AT5G04140.1	182.3	87.2	107.4	51.9	58.8	216.2	143.7	172.5	346.6	44.0	79.5	172.3	59.3	83.5	102.0	107.3
AT2G28000.1	467.1	348.9	489.8	319.0	80.0	53.1	74.5	75.9	72.4	94.3	71.1	57.4	115.0	107.4	112.2	105.1
AT5G20720.1	239.3	218.1	174.4	111.3	49.4	31.5	57.5	46.9	87.6	132.0	58.6	49.2	170.6	155.1	132.7	185.0
AT4G25370.1	330.4	159.9	154.3	170.6	63.5	47.2	53.2	63.5	45.7	81.7	83.7	94.3	115.0	84.8	162.0	201.0
AT5G49910.1	51.3	116.3	136.9	154.3	44.0	52.3	67.8	47.5	96.7	97.5	154.0	160.4	136.7	142.1	95.4	105.5
AT1G42970.1	575.4	773.4	130.8	111.3	98.6	61.1	88.9	70.9	43.8	115.7	102.5	71.0	66.6	88.0	78.6	79.9
AT5G35360.1	239.3	159.9	308.6	267.1	105.9	55.0	125.6	122.8	45.7	106.9	54.4	65.6	66.8	54.3	42.1	9.1
AT4G20360.1	22.8	43.6	154.3	122.4	37.6	37.4	73.4	45.5	87.6	147.8	129.7	143.5	76.0	99.4	84.2	89.1
AT4G17040.1	444.3	232.6	140.9	70.5	28.2	15.7	68.1	46.9	68.6	75.5	92.1	82.0	105.7	88.8	94.4	125.6
AT1G55490.1	246.1	378.0	322.1	151.0	52.0	40.3	73.3	69.1	28.6	103.8	75.3	61.5	63.8	64.2	69.6	59.8
ATCG00670.1	239.3	72.7	73.8	51.9	25.9	9.8	40.4	41.4	34.3	53.4	196.7	45.1	107.6	102.1	90.6	116.5
AT3G26650.1	178.9	234.0	16.8	22.3	86.8	56.4	57.6	53.5	84.9	148.7	94.2	48.4	55.4	33.0	71.3	52.5
AT4G12060.1	239.3	116.3	100.6	70.5	18.8	11.8	25.5	15.2	11.4	66.0	62.8	61.5	100.1	55.7	72.7	121.0
AT1G67090.1	140.1	149.7	92.6	37.1	68.5	89.4	36.6	39.6	73.5	113.8	44.8	73.8	34.3	24.4	38.5	41.6
AT1G09340.1	22.8	14.5	348.9	159.5	32.9	31.5	36.2	51.0	87.6	59.7	58.6	41.0	22.3	26.5	20.4	34.3
AT1G32060.1	216.5	232.6	40.3	74.2	21.2	25.6	28.7	42.8	15.2	81.7	37.7	41.0	22.3	34.5	37.0	20.6
AT2G21330.1	102.5	43.6	67.1	25.6	74.8	42.7	30.4	53.9	28.2	19.2	30.6	39.8	28.7	23.1	34.4	52.1
AT5G38420.1	61.5	46.5	32.2	18.5	36.7	54.1	14.0	18.5	42.7	74.5	31.4	52.1	18.4	11.5	19.3	24.4
AT5G38410.1	61.5	46.5	32.2	18.5	36.7	51.5	14.0	18.5	42.7	74.5	31.4	52.1	18.4	11.5	19.3	24.4
AT3G60750.1	22.8	14.5	0.0	7.4	120.0	151.4	69.2	82.8	160.0	138.3	125.5	127.1	46.4	71.6	85.5	77.6
AT4G38970.1	91.1	0.0	80.5	56.0	87.5	102.8	44.1	64.7	105.1	84.6	74.1	103.8	49.3	61.8	48.5	110.1
AT2G01210.1	353.2	726.8	301.9	70.5	56.5	33.4	26.6	23.5	7.6	15.7	12.6	28.7	31.5	0.0	1.3	43.4
AT4G21280.1	45.6	0.0	10.7	53.0	56.7	43.8	21.9	36.4	3.8	23.6	28.0	30.3	11.1	8.0	12.8	18.7
AT5G38430.1	68.4	46.5	36.9	0.0	39.3	30.9	14.0	17.4	43.4	57.8	26.8	23.0	14.6	6.9	19.9	17.1
AT4G05180.1	11.4	0.0	42.9	106.5	39.8	36.8	14.3	13.2	3.8	14.1	13.8	23.0	3.7	2.7	3.8	13.2
AT1G66670.1	26.2	0.0	40.3	1.1	0.0	20.1	145.8	99.3	72.4	114.8	119.3	69.7	218.8	208.1	191.3	258.1
AT3G12780.1	0.0	0.0	26.8	29.7	63.5	86.9	77.8	89.5	88.7	115.7	124.7	112.8	52.5	61.6	56.3	84.0
AT3G14210.1	0.0	0.0	26.8	26.0	91.7	64.9	38.3	56.6	26.7	59.7	75.3	61.5	109.4	82.2	54.8	57.1
AT4G20850.1	0.0	0.0	6.7	29.7	7.1	13.8	61.7	46.9	110.5	12.6	25.1	41.0	66.8	88.8	98.2	54.8
AT5G35630.1	0.0	43.6	0.0	22.3	23.5	41.3	48.8	52.4	13.7	59.7	67.0	45.1	57.5	55.7	62.5	36.5
AT3G01500.2	0.0	0.0	13.4	29.7	96.9	64.1	45.2	42.5	102.8	57.2	50.2	84.5	13.5	8.7	14.0	29.7
AT3G13470.1	61.5	72.7	147.6	135.8	19.1	22.6	18.2	2.1	28.6	0.0	25.1	0.0	69.7	48.2	68.1	42.0
AT3G63490.1	0.0	0.0	73.8	48.2	18.8	5.9	17.0	17.9	38.1	47.2	20.9	49.2	35.2	26.5	30.6	32.0
AT2G43750.1	0.0	0.0	13.4	33.4	21.2	11.8	22.3	35.9	30.5	37.7	46.0	12.3	33.0	23.9	19.1	18.3
AT3G48870.1; AT5G50920.1	0.00	0.00	15.43	74.19	9.41	14.94	17.24	7.86	21.33	20.75	15.48	8.20	9.83	32.61	36.99	16.67
AT3G11630.1	0.0	0.0	53.7	48.2	26.6	21.0	25.0	23.9	11.4	3.1	58.6	30.8	6.1	1.3	10.2	2.3
AT1G20020.1	0.0	0.0	6.7	3.7	17.6	14.9	16.0	13.8	30.1	3.1	12.6	4.1	5.9	10.7	11.5	18.3
ATCG00780.1	11.4	14.5	26.8	14.8	4.7	5.9	3.2	1.4	0.0	3.1	12.6	8.2	5.6	0.0	5.1	2.3
AT1G36160.1	0.0	0.0	0.0	7.4	735.1	546.5	792.8	923.0	274.2	128.9	83.7	164.0	163.2	385.7	214.7	77.6
AT5G02500.1	0.0	0.0	0.0	9.3	117.9	131.3	78.2	66.9	6.9	47.2	124.7	46.8	130.4	108.7	124.0	164.0
AT1G20620.1	0.0	0.0	0.0	3.7	72.5	81.6	40.6	75.1	50.3	28.3	112.6	144.8	152.2	83.2	77.9	76.3
ATCG00120.1	0.0	0.0	0.0	26.0	51.8	125.8	52.1	22.1	26.7	44.0	108.8	123.0	76.0	63.6	99.5	25.1
AT1G72370.1	0.0	0.0	0.0	1.9	25.9	25.6	53.2	42.8	41.9	34.6	75.3	86.1	66.8	50.4	58.7	66.2
AT3G27925.1	0.0	0.0	53.7	59.3	2.4	0.0	58.5	60.7	41.9	62.9	41.8	28.7	59.3	59.6	47.2	61.7
ATCG00480.1	0.0	0.0	0.0	14.8	105.6	136.6	31.9	19.3	19.0	31.4	62.8	89.0	20.4	19.9	42.1	16.0
AT3G55800.1	22.8	0.0	0.0	0.0	37.6	64.9	41.5	41.4	26.7	34.6	29.3	20.5	40.8	34.5	26.8	43.4
AT2G04030.1	0.0	0.0	0.0	192.9	11.8	11.8	22.0	6.9	26.7	28.3	29.3	12.3	11.1	29.2	35.1	7.5
AT5G30510.1	0.0	0.0	33.5	81.6	0.0	3.9	18.1	12.4	7.6	62.9	25.1	8.2	27.8	38.4	26.8	22.8
AT1G12900.1	178.9	359.1	0.0	0.0	24.2	2.6	27.0	34.4	58.3	21.7	0.0	48.4	5.9	20.9	3.2	4.6
AT5G28540.1	0.0	0.0	2.0	0.0	10.4	14.9	10.0	16.7	1.9	9.7	40.6	49.2	30.2	26.9	20.7	18.0
AT1G06680.1	0.0	0.0	33.5	48.2	28.2	23.6	18.1	16.6	15.2	3.1	20.9	32.8	3.7	0.0	1.3	22.8
AT2G33800.1	0.0	0.0	13.4	22.3	9.4	3.9	9.6	8.3	11.4	3.1	8.4	0.0	5.6	17.2	17.9	18.3
AT5G17920.1	0.0	0.0	0.0	0.0	174.8	278.9	97.8	110.0	199.2	111.9	85.0	77.9	63.2	98.5	67.9	71.7
AT3G55410.1	0.0	0.0	0.0	0.0	11.8	39.3	129.1	81.0	88.0	159.7	108.8	124.7	72.7	93.3	232.2	76.3
AT3G44310.1	0.0	0.0	0.0	0.0	70.6	35.4	56.4	55.2	78.5	91.2	138.1	168.1	170.6	118.0	148.0	155.3
AT5G25980.2	0.0	0.0	0.0	0.0	58.6	68.0	86.5	117.8	30.5	28.3	104.6	94.3	108.3	70.3	111.1	43.8
AT5G26000.1	0.0	0.0	0.0	0.0	106.1	148.2	67.8	86.4	102.8	59.7	87.9	77.9	75.3	57.0	62.4	63.5
AT1G56070.1	0.0	0.0	0.0	0.0	28.2	64.9	106.4	55.2	34.3	12.6	29.3	32.8	26.0	63.6	74.0	93.6
AT1G18080.1	0.0	0.0	0.0	0.0	27.8	13.8	51.3	44.3	76.2	69.2	50.2	24.6	78.8	54.7	57.8	69.7
AT5G14740.2	0.0	0.0	0.0	0.0	67.7	57.8	39.9	75.9	102.1	59.1	62.8	87.8	18.0	21.7	33.2	54.8
AT4G35090.1	0.0	0.0	0.0	0.0	45.2	63.9	38.1	49.3	79.2	59.7	42.3	72.6	58.4	32.1	35.6	28.8
AT3G14067.1	0.0	0.0	0.0	0.0	18.8	15.7	23.4	37.3	19.0	37.7	46.0	28.7	87.2	71.6	53.6	91.3
AT4G33010.1	0.0	0.0	0.0	0.0	57.6	92.4	44.7	67.6	86.8	14.8	37.7	53.3	13.4	17.6	37.4	30.4
AT3G45140.1	79.8	87.2	208.0	593.5	0.0	5.9	1.1	0.0	0.0	0.0	58.6	4.1	16.7	17.2	26.8	9.1
AT2G42600.1	0.0	0.0	0.0	0.0	18.8	60.2	58.1	55.0	27.4	12.6	22.2	45.1	31.5	35.7	47.6	45.7
AT4G37910.1	0.0	0.0	0.0	0.0	25.9	27.3	28.4	18.6	15.2	42.1	67.0	49.2	59.9	46.9	34.8	51.4
AT3G09440.1	0.0	0.0	0.0	0.0	12.9	36.6	41.0	32.0	59.4	36.8	22.2	21.7	35.8	42.9	55.1	18.3
AT2G13360.1	0.0	0.0	0.0	0.0	30.6	35.4	19.2	29.0	91.4	94.3	41.8	61.5	24.1	35.8	30.6	29.7

Supplemental Table 2 (continued)

Accession (experiment # in PPDB: 2164 2165 2190 2191 2192 2193 2224 2225 2208 2209 2210 2211 2214 2215 2216 2217)	10000* NadjSPC exp. 2164 P3 control stroma	10000* NadjSPC exp. 2165 P3 S164A stroma	10000* NadjSPC exp. 2190 P3 control stroma	10000* NadjSPC exp. 2191 P3 S164A stroma	10000* NadjSPC exp. 2192 P3 S164A	10000* NadjSPC exp. 2193 P3 control	10000* NadjSPC exp. 2224 P3 control	10000* NadjSPC exp. 2225 P3 S164A	10000* NadjSPC Exp. 2208 P5 control	10000* NadjSPC exp. 2209 P5 S193A	10000* NadjSPC exp. 2210 P5 control	10000* NadjSPC exp. 2211 P5 S193A	10000* NadjSPC exp. 2214 P5 control	10000* NadjSPC exp. 2215 P5 S193A	10000* NadjSPC exp. 2216 P5 control	10000* NadjSPC exp. 2217 P5 S193A
AT1G13440.1	0.0	0.0	0.0	0.0	64.7	21.6	59.1	19.3	19.0	22.0	23.0	12.3	20.4	23.9	63.1	25.1
AT1G04410.1	0.0	0.0	0.0	0.0	40.0	45.2	28.7	35.9	49.5	47.2	29.3	36.9	22.3	19.9	17.9	25.1
AT4G37930.1	0.0	0.0	0.0	0.0	41.2	51.1	21.3	37.3	68.6	31.4	12.6	61.5	16.7	17.2	20.4	11.4
AT1G09490.1	0.0	0.0	0.0	0.0	4.7	3.9	26.6	19.3	33.5	50.3	41.8	32.8	20.4	35.8	38.3	32.0
AT3G03780.1	0.0	0.0	0.0	0.0	46.3	69.0	22.5	28.0	52.2	13.8	61.5	49.2	5.4	3.6	11.2	26.5
AT1G68010.1	0.0	0.0	0.0	0.0	54.1	72.7	25.5	31.7	11.4	53.4	16.7	32.8	5.6	5.3	17.9	11.4
AT3G52930.1	0.0	0.0	0.0	0.0	47.0	29.9	21.1	31.5	19.0	37.7	33.5	20.5	15.4	20.1	25.1	19.6
AT2G06850.1	0.0	0.0	0.0	0.0	21.2	9.8	28.7	37.3	22.9	15.7	37.7	45.1	33.4	11.9	19.1	38.8
AT5G20290.1	0.0	0.0	0.0	0.0	11.8	2.0	31.9	29.0	15.2	18.9	41.8	36.9	40.8	26.5	19.1	32.0
AT1G01090.1	0.0	0.0	33.5	37.1	11.8	0.0	12.8	26.2	0.0	18.9	16.7	32.8	50.1	29.2	30.6	27.4
AT5G09660.1	0.0	0.0	0.0	0.0	23.5	27.5	17.0	19.3	80.0	37.7	29.3	32.8	11.1	15.9	20.4	41.1
AT5G66190.1	0.0	0.0	0.0	0.0	50.6	36.2	16.0	23.5	38.5	6.3	33.5	20.5	21.9	13.1	24.2	29.7
AT1G23310.1	0.0	0.0	0.0	0.0	32.9	23.6	24.5	37.3	11.4	47.2	18.8	16.4	18.5	13.3	26.8	6.9
AT1G53240.1	0.0	0.0	0.0	0.0	29.2	30.7	36.8	33.3	38.8	8.8	10.5	11.9	11.1	8.7	9.9	24.0
AT4G14880.1	0.0	0.0	0.0	0.0	35.3	13.8	11.7	22.1	45.7	22.0	16.7	16.4	11.1	10.6	8.9	27.4
AT3G14415.1	0.0	0.0	0.0	0.0	18.8	19.5	14.0	23.6	32.0	25.8	12.6	15.2	5.6	13.0	19.6	18.3
AT1G79550.1	0.0	0.0	0.0	0.0	50.1	40.5	16.1	22.2	5.0	23.9	18.4	9.8	7.0	6.1	6.4	11.9
AT2G36530.1	0.0	0.0	0.0	0.0	35.3	15.7	9.6	19.3	11.4	15.7	25.1	49.2	18.5	8.0	15.3	6.9
AT3G04120.1	0.0	0.0	0.0	0.0	10.6	21.6	3.7	19.3	19.0	22.0	23.0	12.3	20.4	23.9	1.9	25.1
AT5G63570.1	0.0	0.0	0.0	0.0	24.5	10.0	10.6	13.8	5.7	11.0	20.9	36.9	6.5	13.3	13.4	11.4
AT4G16155.1	0.0	0.0	127.5	50.4	2.4	0.0	4.3	6.9	0.0	9.4	12.6	16.4	22.3	2.7	11.5	9.1
AT4G31990.1	0.0	0.0	0.0	0.0	19.8	11.8	16.6	19.9	7.6	25.2	29.3	16.4	1.9	1.3	8.9	6.9
AT2G24270.1	0.0	0.0	0.0	0.0	30.6	31.5	5.3	17.9	22.9	9.4	4.2	24.6	1.9	6.6	5.1	9.1
AT4G09000.1	0.0	0.0	0.0	0.0	16.5	14.7	13.0	9.1	11.4	6.3	20.9	16.4	10.4	12.2	1.3	15.1
AT5G17710.1	11.4	0.0	20.1	7.4	11.8	2.0	10.6	16.6	0.0	3.1	0.0	0.0	18.5	10.6	11.5	9.1
AT3G14310.1	0.0	0.0	0.0	0.0	18.8	7.9	9.6	13.8	19.0	12.6	16.7	16.4	5.6	4.0	3.8	9.1
AT3G52380.1	0.0	0.0	33.5	7.4	14.1	3.9	7.4	8.3	0.0	0.0	25.1	8.2	16.7	5.3	1.3	16.0
AT3G47520.1	0.0	0.0	6.7	14.8	7.1	6.7	7.9	6.8	48.8	4.4	0.0	17.6	0.0	5.8	4.3	5.7
AT1G32990.1	11.4	14.5	26.8	18.5	4.7	3.9	7.4	6.9	0.0	0.0	0.0	4.1	1.9	0.0	5.1	4.6
AT1G07320.1	0.0	0.0	13.4	7.4	14.1	2.0	3.2	2.8	3.8	6.3	0.0	0.0	7.4	6.6	3.8	6.9
ATCG00800.1	11.4	0.0	20.1	14.8	0.0	0.0	7.4	5.5	3.8	0.0	8.4	4.1	3.7	2.7	5.1	6.9
AT5G40950.1	0.0	14.5	6.7	3.7	0.0	0.0	5.3	5.5	0.0	12.6	8.4	4.1	11.1	1.3	2.6	6.9
AT5G67360.1	0.0	0.0	0.0	0.0	32.9	27.5	41.5	60.7	0.0	44.0	71.1	73.8	178.0	110.0	52.3	109.6
AT1G03090.2	0.0	0.0	0.0	0.0	58.8	55.0	131.9	147.6	3.8	12.6	4.2	0.0	68.6	43.7	7.7	6.9
AT4G23600.1	0.0	0.0	0.0	0.0	21.2	62.9	8.5	30.4	0.0	6.3	79.5	131.2	81.6	21.2	39.5	16.0
AT3G14420.1	0.0	0.0	0.0	0.0	43.5	37.5	29.6	26.1	74.7	84.6	12.6	29.9	0.0	22.8	26.3	18.3
AT1G11860.1	0.0	0.0	0.0	0.0	30.6	43.2	24.5	33.1	0.0	66.0	33.5	24.6	9.3	13.3	19.1	18.3
AT5G24780.1	0.0	0.0	0.0	0.0	9.4	15.1	16.0	8.8	3.8	0.0	92.1	49.2	40.4	19.9	26.8	13.7
AT3G54050.1	0.0	0.0	0.0	0.0	28.2	29.5	27.7	26.2	19.0	56.6	12.6	32.8	0.0	2.7	11.5	6.9
AT5G09590.1	0.0	0.0	0.0	0.0	0.0	10.0	15.2	15.9	7.6	17.6	20.9	32.8	32.8	15.4	21.3	28.5
AT1G47128.1	0.0	0.0	0.0	0.0	4.7	0.0	13.8	12.4	22.9	28.3	16.7	12.3	26.0	18.6	28.1	32.0
AT2G26080.1	0.0	0.0	0.0	0.0	8.2	57.0	28.7	22.1	16.0	7.2	0.0	8.2	8.9	3.6	11.1	15.3
AT3G55440.1	0.0	0.0	0.0	0.0	53.4	28.7	20.1	21.8	45.7	9.4	4.2	36.9	0.0	1.3	3.8	4.6
AT2G15620.1	0.0	0.0	0.0	7.4	28.2	33.4	12.8	26.2	26.7	3.1	0.0	0.0	7.4	8.0	6.4	18.3
AT3G62030.1	0.0	0.0	13.4	0.0	32.9	41.3	14.9	27.6	11.4	9.4	12.6	8.2	0.0	0.0	3.8	6.9
AT1G23740.1	0.0	0.0	0.0	0.0	23.5	15.7	9.6	12.4	49.5	6.3	0.0	36.9	5.6	6.6	6.4	6.9
AT3G09820.2	0.0	0.0	0.0	0.0	21.2	15.7	3.7	15.9	0.0	9.4	8.4	20.5	13.0	4.0	22.3	4.6
AT4G27520.1	0.0	0.0	0.0	0.0	14.1	25.6	10.6	20.7	15.2	0.0	4.2	16.4	3.7	5.3	5.1	9.1
AT3G15190.1	79.8	130.8	107.4	51.9	7.1	3.9	4.3	5.5	0.0	9.4	4.2	8.2	0.0	0.0	0.0	0.0
AT1G63940.2	0.0	0.0	0.0	0.0	28.2	15.7	12.8	8.3	7.6	12.6	16.7	28.7	3.7	1.3	0.0	6.9
AT3G58610.1	0.0	0.0	0.0	0.0	23.5	23.6	6.4	17.9	19.0	6.3	0.0	4.1	3.7	1.3	5.1	2.3
AT3G17390.1	0.0	0.0	0.0	0.0	7.1	5.9	15.9	5.8	0.0	6.3	4.2	12.3	15.9	4.0	11.5	9.1
AT1G30120.1	0.0	0.0	23.5	11.1	18.8	5.9	3.7	9.7	3.8	0.0	0.0	0.0	10.2	6.6	11.5	11.4
AT2G47390.1	0.0	0.0	0.0	0.0	16.5	15.7	6.4	12.4	22.9	9.4	8.4	12.3	0.0	5.3	2.6	6.9
AT1G05190.1	0.0	0.0	6.7	14.8	9.4	2.0	14.9	9.7	0.0	0.0	12.6	4.1	5.6	1.3	7.7	0.0
AT4G26530.1	0.0	0.0	0.0	0.0	0.0	7.5	2.7	1.7	15.2	25.2	12.6	16.4	5.2	8.6	2.9	7.8
AT1G49240.1	0.0	0.0	0.0	0.0	1.6	0.0	9.5	2.1	5.7	2.5	4.2	12.3	7.0	9.1	10.7	3.7
AT3G18780.2	0.0	0.0	0.0	0.0	1.6	0.0	9.5	2.1	5.7	2.5	4.2	12.3	7.0	9.1	10.7	3.7
AT1G16880.1	0.0	0.0	0.0	14.8	11.8	0.0	6.4	0.0	19.0	3.1	8.4	4.1	7.4	2.7	1.3	16.0
AT4G28750.1	0.0	0.0	26.8	11.1	7.1	3.9	2.1	4.1	0.0	6.3	12.6	8.2	1.9	0.0	2.6	0.0
AT2G40840.1	0.0	0.0	0.0	0.0	7.1	7.9	47.9	17.9	3.8	12.6	0.0	0.0	22.3	47.7	43.4	36.5
AT3G19170.1	0.0	0.0	0.0	0.0	32.9	62.9	37.2	37.3	60.9	15.7	0.0	0.0	5.6	8.0	21.7	27.4
AT1G78900.1	0.0	0.0	0.0	0.0	9.4	27.5	19.2	17.9	22.9	12.6	0.0	0.0	50.1	27.8	39.5	34.3
AT3G09840.1	0.0	0.0	0.0	0.0	0.0	5.9	35.1	5.1	34.3	40.9	16.7	0.0	27.8	54.3	15.3	27.4
AT3G23990.1	0.0	0.0	0.0	0.0	27.5	29.5	26.0	16.3	7.6	0.0	8.4	0.0	29.3	36.8	22.1	20.6
AT2G05920.1	0.0	0.0	0.0	0.0	0.0	3.9	20.2	22.1	0.0	15.7	4.2	12.3	55.6	25.2	17.9	27.4
AT1G62750.1	0.0	0.0	0.0	0.0	21.2	70.8	22.3	34.5	34.3	9.4	8.4	8.2	0.0	0.0	6.4	6.9
AT2G21170.1	0.0	14.5	0.0	0.0	78.3	20.4	24.6	32.0	26.7	3.1	0.0	0.0	5.6	2.7	0.0	4.6
AT1G07930.1	0.0	0.0	0.0	0.0	18.6	16.3	23.8	13.4	0.0	2.5	0.0	5.3	26.1	19.5	18.9	15.1
AT1G07940.1	0.0	0.0	0.0	0.0	18.6	16.3	23.8	13.4	0.0	2.5	0.0	5.3	26.1	19.5	18.9	15.1
AT1G07920.1	0.0	0.0	0.0	0.0	18.6	16.3	23.8	13.4	0.0	2.5	0.0	5.3	26.1	19.5	18.9	15.1
AT5G60390.1	0.0	0.0	0.0	0.0	18.6	16.3	23.8	13.4	0.0	2.5	0.0	5.3	26.1	19.5	18.9	15.1
AT5G09810.1	0.0	0.0	0.0	0.0	1.6	4.9	14.7	8.3	0.0	23.0	8.4	0.0	23.0	26.2	9.8	40.9
AT3G11830.1	0.0	0.0	0.0	0.0	11.8	5.9	6.4	8.3	3.8	9.4	0.0	0.0	22.3	33.1	24.2	11.4
AT1G56190.1	0.0	0.0	0.0	0.0	17.9	20.2	8.3	22.1	5.0	36.5	15.9	16.8	0.0	7.8	16.5	0.0
AT4G13930.1	0.0	0.0	0.0	0.0	32.9	17.7	8.5	23.5	0.0	0.0	8.4	24.6	1.9	10.6	5.1	9.1
AT3G48730.1	0.0	0.0	0.0	0.0	34.3	25.4	10.6	13.8	5.7	11.0	0.0	4.1	6.5	0.0	13.4	11.4
AT5G04590.1	0.0	0.0	0.0	29.7	30.6	41.3	5.3	9.7	3.8	3.1	0.0	0.0	0.0	1.3	5.1	2.3

Supplemental Table 2 (continued)

Accession (experiment # in PPDB: 2164 2165 2190 2191 2192 2193 2224 2225 2208 2209 2210 2211 2214 2215 2216 2217)	10000* NadjSP C exp. 2164 P3 control stroma	10000* NadjSP C exp. 2165 P3 S164A stroma	10000* NadjSP C exp. 2190 P3 control stroma	10000* NadjSP C exp. 2191 P3 S164A stroma	10000* NadjSP C exp. 2192 P3 S164A	10000* NadjSP C exp. 2193 P3 control	10000* NadjSP C exp. 2224 P3 control	10000* NadjSP C exp. 2225 P3 S164A	10000* NadjSP C Exp. 2208 P5 control	10000* NadjSP C exp. 2209 P5 S193A	10000* NadjSP C exp. 2210 P5 control	10000* NadjSP C exp. 2211 P5 S193A	10000* NadjSP C exp. 2214 P5 control	10000* NadjSP C exp. 2215 P5 S193A	10000* NadjSP C exp. 2216 P5 control	10000* NadjSP C exp. 2217 P5 S193A
AT1G03475.1	0.0	0.0	0.0	0.0	11.8	7.9	9.6	23.5	26.7	0.0	8.4	20.5	3.7	6.6	6.4	0.0
AT1G43670.1	0.0	0.0	0.0	0.0	23.5	21.6	5.3	8.3	15.2	0.0	8.4	8.2	0.0	8.0	8.9	9.1
AT3G25860.1	0.0	0.0	80.5	82.3	11.8	3.9	2.1	0.0	0.0	0.0	4.2	8.2	7.4	2.7	3.8	0.0
AT1G53310.1	0.0	0.0	0.0	0.0	4.7	10.6	17.5	7.0	6.9	12.6	11.3	0.0	0.0	6.8	8.5	11.4
AT5G19510.1	0.0	0.0	0.0	0.0	18.8	17.1	9.6	8.3	7.6	0.0	12.6	12.3	5.6	8.0	0.0	11.4
AT1G68560.1	0.0	0.0	0.0	0.0	4.7	13.8	5.3	11.0	15.2	12.6	0.0	0.0	11.1	9.3	6.4	11.4
AT2G31610.1	0.0	0.0	0.0	0.0	0.0	1.4	18.3	5.1	7.6	3.1	0.0	10.7	5.6	1.7	7.7	22.8
AT2G34590.1	0.0	0.0	23.5	11.1	0.0	5.9	3.7	9.7	3.8	0.0	0.0	0.0	10.2	6.6	11.5	11.4
AT3G22890.1	0.0	0.0	8.7	29.7	7.1	5.9	2.1	11.0	0.0	0.0	0.0	20.5	5.6	4.0	3.8	0.0
AT1G74970.1	0.0	0.0	33.5	29.7	4.7	0.0	2.1	4.1	0.0	6.3	8.4	12.3	1.9	0.0	8.9	0.0
AT1G69740.1	0.0	0.0	26.8	11.1	7.1	2.0	7.4	8.3	0.0	12.6	0.0	8.2	0.0	1.3	3.8	0.0
AT1G35680.1	11.4	14.5	47.0	18.5	4.7	3.9	5.3	2.8	0.0	0.0	0.0	16.4	0.0	0.0	0.0	2.3
AT5G65220.1	11.4	0.0	13.4	14.8	0.0	0.0	5.3	6.9	0.0	9.4	4.2	8.2	0.0	0.0	1.3	11.4
AT3G12580.1	0.0	0.0	2.0	9.3	0.0	0.0	1.3	0.6	6.9	0.0	5.0	8.6	0.0	1.3	3.8	4.1
AT5G65750.1	0.0	0.0	0.0	0.0	0.0	0.0	54.1	21.1	11.0	57.2	0.0	23.0	18.2	47.2	60.0	17.4
AT3G18490.1	0.0	0.0	0.0	0.0	0.0	0.0	18.1	15.2	0.0	37.7	58.6	32.8	37.1	38.4	26.8	29.7
AT4G13940.1	0.0	0.0	0.0	0.0	25.9	19.7	29.8	22.1	0.0	0.0	0.0	12.3	22.3	21.2	21.0	13.7
AT1G65930.1	0.0	0.0	0.0	0.0	30.6	31.5	27.7	29.0	26.7	25.2	0.0	16.4	0.0	11.9	6.4	0.0
AT5G42020.1	0.0	0.0	2.0	0.0	0.0	0.0	14.9	0.0	1.9	39.0	40.6	0.0	35.6	30.0	13.8	22.6
AT1G19570.1	0.0	0.0	0.0	0.0	32.9	23.6	12.8	22.1	0.0	0.0	41.8	90.2	5.6	0.0	2.6	9.1
AT2G24200.1	0.0	0.0	0.0	0.0	37.6	49.1	5.3	19.3	0.0	3.1	16.7	0.0	7.4	13.3	5.1	0.0
AT2G13440.1	0.0	0.0	0.0	7.4	0.0	0.0	6.4	2.8	0.0	6.3	4.2	0.0	16.7	30.5	31.9	16.0
AT4G35830.1	0.0	0.0	0.0	0.0	22.8	38.5	10.7	8.7	14.5	3.1	0.0	0.0	0.0	3.3	8.9	20.3
AT3G25520.1	0.0	0.0	0.0	0.0	0.0	1.0	23.9	6.2	15.2	7.9	0.0	0.0	17.6	18.3	12.8	3.4
AT5G56010.1	0.0	0.0	0.0	0.0	0.0	0.0	12.2	0.0	22.9	8.5	11.3	8.2	10.0	26.2	20.7	5.3
AT5G56030.1	0.0	0.0	0.0	0.0	0.0	0.0	12.2	0.0	22.9	8.5	11.3	8.2	10.0	26.2	20.7	5.3
AT1G66970.1	0.0	0.0	0.0	0.0	0.0	18.7	9.6	14.5	68.6	0.0	8.4	16.4	0.0	2.7	6.4	4.6
AT1G21680.1	0.0	0.0	0.0	0.0	0.0	2.0	4.3	6.9	7.6	22.0	0.0	0.0	22.3	13.3	6.4	29.7
AT5G19220.1	0.0	0.0	93.9	85.3	0.0	5.9	3.2	6.9	0.0	6.3	0.0	0.0	1.9	4.0	3.8	0.0
AT1G11840.1	0.0	0.0	0.0	0.0	19.5	17.1	13.8	18.9	11.4	0.0	8.4	16.4	1.9	0.0	0.0	4.6
AT4G14960.2	0.0	0.0	0.0	0.0	0.0	0.0	7.6	0.8	0.0	9.4	3.3	32.8	8.5	31.0	9.4	1.4
AT3G11770.1	0.0	0.0	0.0	0.0	4.7	2.0	13.8	15.2	0.0	0.0	29.3	0.0	11.1	6.6	8.9	4.6
AT3G63140.1	0.0	0.0	93.9	37.1	9.4	0.0	11.7	11.0	7.6	0.0	0.0	0.0	1.9	2.7	0.0	2.3
AT5G39740.1	0.0	0.0	0.0	0.0	0.0	1.0	2.7	6.2	15.2	7.9	0.0	0.0	17.6	10.9	12.8	24.0
AT4G09320.1	0.0	0.0	0.0	0.0	28.2	17.7	9.6	5.5	0.0	15.7	12.6	24.6	3.7	0.0	0.0	4.6
AT5G14200.1	0.0	0.0	0.0	0.0	21.2	29.5	6.4	6.9	0.0	3.1	16.7	28.7	0.0	0.0	0.9	4.6
AT5G03350.1	0.0	0.0	0.0	0.0	2.4	5.9	12.8	16.6	0.0	0.0	12.6	8.2	11.1	2.7	0.0	18.3
AT3G52880.1	0.0	0.0	0.0	0.0	11.8	11.8	8.5	17.9	0.0	15.7	12.6	16.4	0.0	5.3	1.3	0.0
AT3G14240.1	0.0	0.0	0.0	0.0	0.0	0.0	8.5	9.7	0.0	12.6	8.4	16.4	13.0	2.7	8.9	16.0
AT5G16390.1	0.0	0.0	0.0	11.1	18.8	7.9	14.9	8.3	0.0	0.0	0.0	8.2	5.6	2.7	3.8	0.0
AT5G36700.1	0.0	0.0	0.0	0.0	16.5	6.9	9.6	9.7	20.9	9.4	10.5	16.4	0.0	0.0	0.0	6.9
AT5G36790.1	0.0	0.0	0.0	0.0	16.5	6.9	9.6	9.7	20.9	9.4	10.5	16.4	0.0	0.0	0.0	6.9
AT4G14040.1	0.0	0.0	0.0	0.0	16.5	21.6	6.0	5.5	0.0	0.0	4.2	16.4	1.9	4.0	2.6	0.0
AT2G18020.1	0.0	0.0	0.0	0.0	0.0	0.0	9.6	2.8	15.2	15.1	16.7	0.0	5.6	6.6	6.4	2.3
AT2G32080.1	0.0	0.0	0.0	0.0	0.0	0.0	4.3	1.4	15.2	18.9	4.2	8.2	5.6	18.6	0.0	2.3
AT1G30230.1	0.0	0.0	0.0	0.0	14.1	6.5	5.3	2.8	15.2	0.0	0.0	0.0	3.7	5.3	5.1	11.4
AT2G40490.1	0.0	0.0	40.3	48.2	4.7	7.9	3.2	1.4	0.0	0.0	0.0	4.1	0.0	1.3	1.3	0.0
AT4G16760.1	0.0	0.0	0.0	0.0	0.0	3.9	5.3	2.8	7.6	12.6	0.0	0.0	9.3	5.3	6.4	2.3
AT5G50850.1	0.0	0.0	0.0	0.0	4.7	0.0	8.5	1.4	7.6	6.3	0.0	8.2	3.7	5.3	5.1	0.0
AT2G37220.1	0.0	0.0	0.0	0.0	7.1	2.0	4.3	11.0	0.0	0.0	8.4	12.3	3.7	2.7	0.0	2.3
ATCG00830.1	0.0	0.0	0.0	3.7	2.4	1.0	5.9	2.8	0.0	0.0	0.0	0.0	3.7	2.7	10.8	3.4
ATCG01310.1	0.0	0.0	0.0	3.7	2.4	1.0	5.9	2.8	0.0	0.0	0.0	0.0	3.7	2.7	10.8	3.4
AT5G20020.1	0.0	0.0	0.0	0.0	5.4	3.9	4.7	6.5	3.8	3.1	0.0	7.0	8.7	2.3	0.0	0.0
AT5G20010.1	0.0	0.0	0.0	0.0	5.4	3.9	4.7	6.5	3.8	3.1	0.0	7.0	8.7	2.3	0.0	0.0
AT5G55190.1	0.0	0.0	0.0	0.0	5.4	3.9	4.7	6.5	3.8	3.1	0.0	7.0	8.7	2.3	0.0	0.0
AT3G53870.1	0.0	0.0	0.0	0.0	0.0	1.4	2.7	5.1	7.6	3.1	0.0	10.7	5.6	1.7	7.7	0.0
AT3G27830.1	28.5	43.6	36.9	13.0	2.4	2.9	2.1	0.0	0.0	4.7	8.4	0.0	0.0	0.0	0.0	0.0
AT3G27850.1	28.5	43.6	36.9	13.0	2.4	2.9	2.1	0.0	0.0	4.7	8.4	0.0	0.0	0.0	0.0	0.0
AT4G12420.1	0.0	0.0	0.0	0.0	2.4	9.8	2.1	2.8	22.9	3.1	4.2	8.2	0.0	0.0	0.0	2.3
AT4G27440.1	11.4	0.0	0.0	7.4	0.0	2.0	4.3	1.4	0.0	0.0	0.0	0.0	3.7	2.7	6.4	4.6
AT5G02490.1	0.0	0.0	0.0	0.0	0.0	0.0	1.3	0.6	6.9	8.5	5.0	8.6	0.0	6.1	4.6	4.1
AT2G20890.1	0.0	0.0	0.0	0.0	2.4	3.9	3.2	4.1	3.8	0.0	12.6	0.0	5.6	1.3	0.0	4.6
AT4G13430.1	0.0	0.0	0.0	0.0	2.4	3.9	3.2	4.1	0.0	0.0	0.0	4.1	1.9	2.7	2.6	4.6
AT1G52000.1	0.0	0.0	0.0	0.0	16.5	37.4	16.0	38.6	0.0	0.0	8.4	0.0	29.7	18.6	10.2	0.0
AT3G03960.1	0.0	0.0	0.0	0.0	0.0	5.9	14.9	2.8	0.0	0.0	4.2	0.0	44.5	35.8	20.4	45.7
AT2G32730.1	0.0	0.0	0.0	0.0	0.0	0.0	18.1	2.8	5.7	9.4	0.0	0.0	4.6	22.5	44.8	11.4
AT2G20580.1	0.0	0.0	0.0	0.0	0.0	3.9	18.1	4.1	0.0	9.4	0.0	0.0	16.7	21.2	20.4	25.1
AT1G63770.3	0.0	0.0	0.0	0.0	11.8	55.0	21.3	13.8	15.2	0.0	0.0	0.0	0.0	1.3	1.3	6.9
AT4G29060.1	0.0	0.0	73.8	103.9	16.5	17.7	7.4	0.0	7.6	0.0	0.0	0.0	0.0	2.7	6.4	0.0
AT5G11720.1	0.0	0.0	0.0	0.0	0.0	0.0	4.3	9.7	0.0	25.2	16.7	0.0	14.8	11.9	11.5	22.8
AT5G56500.1	0.0	0.0	0.0	95.3	4.2	15.5	0.0	4.7	0.0	0.0	0.0	0.0	11.1	4.4	6.3	3.0
AT5G08280.1	0.0	0.0	0.0	0.0	16.5	17.7	10.6	8.3	41.9	0.0	0.0	12.3	0.0	2.7	1.3	0.0
AT4G23670.1	0.0	0.0	0.0	0.0	30.6	25.6	4.3	13.8	7.6	15.7	4.2	4.1	0.0	0.0	0.0	0.0
AT4G02520.1	0.0	0.0	0.0	0.0	11.8	15.7	7.4	15.2	15.2	0.0	29.3	16.4	0.0	1.3	0.0	0.0
AT4G22010.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	4.1	0.0	6.3	4.2	4.1	5.6	27.8	16.6	2.3
AT2G38230.1	0.0	0.0	0.0	0.0	0.0	2.6	10.1	6.8	11.4	0.0	0.0	0.0	7.4	8.0	8.4	16.9
AT5G48300.1	0.0	0.0	40.3	63.1	0.0	7.9	4.3	1.4	0.0	3.1	0.0	0.0	3.7	0.0	7.7	0.0
AT4G33090.1	0.0	0.0	0.0	0.0	9.4	43.2	6.4	5.5	7.6	3.1	4.2	0.0	0.0	0.0	0.0	2.3

Supplemental Table 2 (continued)

Accession (experiment # in PPDB: 2164 2165 2190 2191 2192 2193 2224 2225 2208 2209 2210 2211 2214 2215 2216 2217)	10000* NadjSP C exp. 2164 P3 control stroma	10000* NadjSP C exp. 2165 P3 S164A stroma	10000* NadjSP C exp. 2190 P3 control stroma	10000* NadjSP C exp. 2191 P3 S164A stroma	10000* NadjSP C exp. 2192 P3 S164A	10000* NadjSP C exp. 2193 P3 control	10000* NadjSP C exp. 2224 P3 control	10000* NadjSP C exp. 2225 P3 S164A	10000* NadjSP C Exp. 2208 P5 control	10000* NadjSP C exp. 2209 P5 S193A	10000* NadjSP C exp. 2210 P5 control	10000* NadjSP C exp. 2211 P5 S193A	10000* NadjSP C exp. 2214 P5 control	10000* NadjSP C exp. 2215 P5 S193A	10000* NadjSP C exp. 2216 P5 control	10000* NadjSP C exp. 2217 P5 S193A
AT4G33680.1	0.0	0.0	0.0	0.0	11.8	7.9	8.5	12.4	0.0	15.7	0.0	12.3	1.9	0.0	2.6	0.0
AT1G29670.1	0.0	0.0	0.0	0.0	4.7	7.9	9.6	9.7	19.0	0.0	0.0	8.2	0.0	1.3	0.0	16.0
AT1G23190.1	0.0	0.0	0.0	0.0	14.1	20.1	4.7	5.9	0.0	3.1	0.0	16.4	0.0	5.3	3.8	0.0
AT5G11670.1	0.0	0.0	0.0	0.0	11.8	21.6	2.1	2.8	19.0	0.0	0.0	0.0	0.0	5.3	2.6	4.6
AT1G22300.1	0.0	0.0	0.0	0.0	18.6	6.5	3.4	7.2	0.0	0.0	4.2	0.0	10.0	5.3	0.0	4.6
AT3G53180.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.4	0.0	6.3	8.4	8.2	9.3	11.9	8.9	9.1
AT2G01250.1	0.0	0.0	0.0	0.0	0.0	0.0	6.7	2.8	22.9	15.7	0.0	16.4	6.5	0.0	1.3	9.1
AT3G02230.1	0.0	0.0	0.0	0.0	4.0	2.6	4.3	4.1	3.8	0.0	0.0	0.0	3.2	17.2	7.4	0.0
AT3G44890.1	0.0	0.0	0.0	0.0	0.0	0.0	7.4	11.0	0.0	3.1	8.4	12.3	3.7	0.0	6.4	2.3
AT4G39260.1	0.0	0.0	0.0	11.1	10.6	13.8	7.4	4.8	0.0	0.0	0.0	0.0	1.9	0.0	2.6	2.3
AT5G43940.1	0.0	0.0	0.0	0.0	11.8	5.9	7.4	5.5	0.0	6.3	4.2	12.3	0.0	0.0	3.8	0.0
AT2G21660.1	0.0	0.0	0.0	0.0	10.6	0.0	7.4	13.1	0.0	0.0	4.2	4.1	3.7	2.7	1.3	0.0
AT3G18190.1	0.0	0.0	0.0	0.0	2.4	0.0	3.2	5.5	3.8	0.0	0.0	0.0	5.6	9.3	8.9	2.3
AT5G58330.1	0.0	0.0	0.0	26.0	0.0	0.0	2.1	4.1	0.0	6.3	0.0	16.4	0.0	2.7	6.4	2.3
AT3G54400.1	0.0	0.0	0.0	0.0	9.4	11.8	4.3	6.9	0.0	6.3	4.2	12.3	1.9	0.0	0.0	0.0
AT1G50010.1	0.0	0.0	0.0	0.0	0.0	0.0	4.8	0.8	0.0	6.6	3.3	0.0	8.5	6.5	9.4	1.4
AT1G04820.1	0.0	0.0	0.0	0.0	0.0	0.0	4.8	0.8	0.0	6.6	3.3	0.0	8.5	6.5	9.4	1.4
AT5G19770.1	0.0	0.0	0.0	0.0	0.0	0.0	5.3	0.8	0.0	2.5	3.3	0.0	3.0	8.5	12.0	1.4
AT5G19780.1	0.0	0.0	0.0	0.0	0.0	0.0	5.3	0.8	0.0	2.5	3.3	0.0	3.0	8.5	12.0	1.4
AT5G15450.1	0.0	0.0	0.0	7.4	0.0	3.9	9.6	0.0	7.6	0.0	0.0	0.0	3.7	4.0	2.6	6.9
AT5G38480.1	0.0	0.0	0.0	0.0	0.0	2.4	7.4	2.8	3.8	3.1	8.4	0.0	13.0	4.5	0.0	0.0
AT3G04790.1	0.0	0.0	0.0	0.0	4.7	7.9	3.2	5.5	19.0	0.0	8.4	8.2	1.9	0.0	0.0	0.0
AT2G30110.1	0.0	0.0	0.0	0.0	2.4	2.0	9.6	2.8	0.0	0.0	0.0	8.2	0.0	2.7	6.4	2.3
AT1G66200.1	0.0	0.0	0.0	0.0	7.1	0.0	2.2	8.3	3.8	12.6	0.0	0.0	1.9	2.7	2.6	0.0
AT2G43030.1	0.0	0.0	0.0	14.8	0.0	0.0	5.3	2.8	3.8	3.1	0.0	4.1	11.1	0.0	0.0	2.3
AT1G79850.1	0.0	14.5	13.4	18.5	4.7	0.0	0.0	4.1	0.0	6.3	0.0	8.2	0.0	0.0	0.0	6.9
AT1G03680.1	0.0	0.0	0.0	0.0	4.7	11.8	3.2	2.8	0.0	3.1	8.4	8.2	0.0	0.0	0.0	4.6
AT5G03300.1	0.0	0.0	0.0	0.0	2.4	0.0	3.7	2.1	0.0	9.4	8.4	0.0	0.0	4.0	1.9	4.6
AT5G20630.1	0.0	0.0	0.0	0.0	9.4	3.9	1.1	4.1	7.6	0.0	8.4	4.1	1.9	0.0	0.0	0.0
AT3G11130.1	0.0	0.0	0.0	0.0	0.0	27.5	21.3	0.0	91.4	0.0	0.0	18.5	3.7	27.4	39.3	0.0
AT3G20050.1	0.0	0.0	0.0	0.0	2.4	0.0	3.2	0.0	0.0	0.0	0.0	4.1	31.5	27.8	17.9	6.9
AT2G30860.1	0.0	0.0	0.0	0.0	16.5	23.6	11.7	19.3	19.0	0.0	12.6	8.2	0.0	0.0	0.0	0.0
AT3G04840.1	0.0	0.0	0.0	0.0	0.0	0.0	14.7	2.3	29.3	12.6	0.0	0.0	0.0	5.6	17.2	16.0
AT2G33210.1	0.0	0.0	0.0	0.0	7.8	11.8	4.9	14.1	0.0	0.0	0.0	0.0	5.9	13.5	15.1	0.0
AT4G26970.1	0.0	0.0	0.0	0.0	17.2	17.1	15.4	15.6	0.0	3.1	0.0	0.0	0.0	1.3	0.0	11.6
AT4G01310.1	0.0	0.0	0.0	44.5	7.1	5.9	7.4	23.5	0.0	0.0	0.0	0.0	5.6	0.0	0.0	6.9
AT5G26360.1	0.0	0.0	0.0	0.0	0.0	0.0	8.5	2.8	0.0	0.0	4.2	0.0	16.7	17.2	15.3	6.9
AT3G09200.1	0.0	0.0	0.0	0.0	0.0	0.0	4.3	0.0	30.5	12.6	0.0	4.1	5.6	15.9	19.1	0.0
AT2G47940.1	0.0	0.0	0.0	29.7	0.0	0.0	4.3	0.0	0.0	0.0	0.0	4.1	18.5	9.3	11.5	11.4
AT4G08870.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	8.3	0.0	0.0	12.6	28.7	14.8	14.6	7.7	6.9
AT5G02240.1	0.0	0.0	0.0	0.0	16.5	23.6	10.6	13.8	3.8	0.0	4.2	0.0	0.0	0.0	0.0	6.9
AT5G53460.1	0.0	0.0	0.0	0.0	0.0	2.0	17.0	9.7	49.5	3.1	0.0	4.1	0.0	0.0	5.1	0.0
AT2G05710.1	0.0	0.0	0.0	0.0	14.3	14.9	17.5	3.3	24.0	3.1	0.0	0.0	0.0	3.3	0.0	0.0
AT5G15650.1	0.0	0.0	0.0	0.0	4.0	2.6	0.0	0.0	3.8	0.0	0.0	4.1	19.1	17.2	16.8	0.0
AT1G57720.1	0.0	0.0	0.0	0.0	26.8	11.8	10.2	1.4	0.0	6.3	0.0	0.0	4.8	9.9	0.0	0.0
AT2G45470.1	0.0	0.0	0.0	0.0	4.7	22.6	6.4	15.2	11.4	3.1	0.0	0.0	0.0	0.0	3.8	0.0
AT3G13920.1	0.0	0.0	0.0	0.0	0.0	0.0	4.5	1.8	0.0	4.1	4.2	0.0	4.3	10.2	24.9	0.0
AT5G10450.3	0.0	0.0	0.0	0.0	6.8	3.9	9.8	12.4	0.0	0.0	0.0	0.0	6.9	10.1	0.0	2.3
AT5G01410.1	0.0	0.0	0.0	0.0	9.4	2.6	10.1	11.2	0.0	0.0	0.0	0.0	1.9	0.0	5.6	12.8
AT3G15356.1	0.0	0.0	0.0	7.4	4.7	9.4	10.6	5.5	22.9	0.0	0.0	0.0	7.4	0.0	0.0	0.0
AT1G72150.1	0.0	0.0	0.0	0.0	0.0	3.9	13.8	2.8	0.0	0.0	8.4	0.0	2.8	8.0	6.4	0.0
AT1G01080.1	0.0	0.0	67.1	26.0	0.0	0.0	4.3	1.4	0.0	0.0	16.7	0.0	0.0	4.0	0.0	4.6
AT2G44650.1	0.0	0.0	0.0	3.7	4.7	2.0	0.0	0.0	0.0	0.0	0.0	0.0	7.4	4.0	10.2	27.4
AT1G11910.1	0.0	0.0	0.0	0.0	0.0	0.0	4.3	4.1	15.2	0.0	8.4	0.0	0.0	10.6	7.7	9.1
AT2G20420.1	0.0	0.0	0.0	0.0	11.8	7.9	9.6	9.7	0.0	0.0	8.4	4.1	0.0	0.0	2.6	0.0
AT2G28190.1	0.0	0.0	0.0	0.0	16.5	7.9	5.3	6.9	0.0	0.0	8.4	0.0	5.6	0.0	0.0	6.9
AT1G35160.1	0.0	0.0	0.0	0.0	8.2	0.0	5.3	10.6	7.6	0.0	0.0	0.0	5.2	4.5	0.0	7.8
AT2G17360.1	0.0	0.0	0.0	0.0	0.0	0.0	11.4	0.0	0.0	8.5	5.4	0.0	6.1	3.6	6.0	5.3
AT5G07090.1	0.0	0.0	0.0	0.0	0.0	0.0	11.4	0.0	0.0	8.5	5.4	0.0	6.1	3.6	6.0	5.3
AT2G30970.1	0.0	0.0	0.0	0.0	11.8	5.9	5.3	11.0	0.0	0.0	0.0	0.0	0.0	2.7	2.6	4.6
AT5G58290.1	0.0	0.0	0.0	0.0	2.4	0.0	2.9	4.6	0.0	12.6	0.0	0.0	3.7	4.0	13.4	0.0
ATCG00900.1	0.0	29.1	70.5	26.0	2.4	2.0	2.1	0.0	0.0	6.3	0.0	0.0	0.0	0.0	0.0	0.0
ATCG01240.1	0.0	29.1	70.5	26.0	2.4	2.0	2.1	0.0	0.0	6.3	0.0	0.0	0.0	0.0	0.0	0.0
AT5G06290.1	0.0	0.0	0.0	0.0	8.9	8.5	5.9	7.9	7.6	0.0	0.0	10.3	3.2	0.0	0.0	0.0
AT2G41840.1	0.0	0.0	0.0	0.0	0.0	0.0	7.4	2.8	3.8	0.0	0.0	0.0	7.4	0.4	11.5	4.6
AT3G23810.1	0.0	0.0	0.0	0.0	0.0	2.0	6.4	4.1	0.0	0.0	0.0	12.3	5.6	8.0	1.9	0.0
AT3G25530.1	0.0	0.0	0.0	0.0	2.4	3.9	8.5	9.7	3.8	0.0	0.0	0.0	5.6	0.0	0.0	2.3
AT5G65620.1	0.0	0.0	0.0	0.0	4.7	11.2	9.6	3.4	0.0	3.1	0.0	0.0	0.0	0.0	1.3	2.3
AT1G10760.1	0.0	0.0	0.0	1.9	0.0	1.0	3.2	1.4	45.7	0.0	0.0	4.1	0.0	0.0	3.8	0.0
AT1G09750.1	0.0	0.0	0.0	0.0	4.7	11.8	0.0	2.8	3.8	9.4	0.0	20.5	0.0	2.7	0.0	0.0
AT3G57260.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.8	30.5	3.1	25.1	4.1	1.9	2.7	0.0	0.0
AT1G78630.1	0.0	0.0	20.1	0.0	4.7	7.9	3.2	2.8	0.0	0.0	8.4	0.0	5.6	0.0	0.0	0.0
AT3G01480.1	0.0	0.0	0.0	0.0	4.7	0.0	6.4	6.9	7.6	3.1	0.0	4.1	0.0	0.0	2.6	0.0
AT5G58420.1	0.0	0.0	0.0	0.0	0.0	0.0	1.8	0.0	0.0	8.5	5.4	0.0	6.1	3.6	6.0	5.3
AT5G35530.1	0.0	0.0	0.0	0.0	0.0	1.4	7.8	5.1	7.6	3.1	0.0	10.7	0.0	1.7	0.0	0.0
AT5G56000.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	8.5	11.3	8.2	10.0	1.7	2.2	5.3
AT5G06450.1	0.0	0.0	0.0	0.0	0.0	0.0	2.1	5.5	0.0	6.3	8.4	0.0	5.6	1.3	0.0	2.3
AT1G14810.1	0.0	0.0	6.7	3.7	14.1	2.0	2.1	1.4	0.0	0.0	0.0	0.0	3.7	0.0	0.0	0.0

Supplemental Table 2 (continued)

Accession (experiment # in PPDB: 2164 2165 2190 2191 2192 2193 2224 2225 2208 2209 2210 2211 2214 2215 2216 2217)	10000* NadjSP C exp. 2164 P3 control stroma	10000* NadjSP C exp. 2165 P3 S164A stroma	10000* NadjSP C exp. 2190 P3 control stroma	10000* NadjSP C exp. 2191 P3 S164A stroma	10000* NadjSP C exp. 2192 P3 S164A	10000* NadjSP C exp. 2193 P3 control	10000* NadjSP C exp. 2224 P3 control	10000* NadjSP C exp. 2225 P3 S164A	10000* NadjSP C Exp. 2208 P5 control	10000* NadjSP C exp. 2209 P5 S193A	10000* NadjSP C exp. 2210 P5 control	10000* NadjSP C exp. 2211 P5 S193A	10000* NadjSP C exp. 2214 P5 control	10000* NadjSP C exp. 2215 P5 S193A	10000* NadjSP C exp. 2216 P5 control	10000* NadjSP C exp. 2217 P5 S193A
AT1G22780.1	0.0	0.0	0.0	0.0	0.0	2.0	3.9	1.8	0.0	0.0	0.0	2.9	3.2	2.3	4.2	0.0
AT1G34030.1	0.0	0.0	0.0	0.0	0.0	2.0	3.9	1.8	0.0	0.0	0.0	2.9	3.2	2.3	4.2	0.0
AT4G09800.1	0.0	0.0	0.0	0.0	0.0	2.0	3.9	1.8	0.0	0.0	0.0	2.9	3.2	2.3	4.2	0.0
AT5G13510.1	0.0	14.5	0.0	3.7	2.4	0.0	3.2	4.1	0.0	0.0	8.4	0.0	0.0	0.0	1.3	0.0
AT5G51970.1	0.0	0.0	0.0	0.0	4.7	0.0	3.2	4.1	0.0	0.0	4.2	4.1	1.9	0.0	0.0	2.3
AT1G47250.1	0.0	0.0	0.0	0.0	3.5	0.0	0.0	2.8	3.8	0.0	6.3	0.0	2.8	2.0	3.2	0.0
AT5G13120.2	0.0	0.0	0.0	0.0	0.0	5.9	1.1	1.4	0.0	0.0	8.4	4.1	3.7	0.0	1.3	0.0
AT2G37190.1	0.0	0.0	0.0	0.0	1.6	2.0	2.1	3.2	0.0	0.0	0.0	5.3	3.7	0.0	0.0	3.0
AT3G53430.1	0.0	0.0	0.0	0.0	1.6	2.0	2.1	3.2	0.0	0.0	0.0	5.3	3.7	0.0	0.0	3.0
AT5G60670.1	0.0	0.0	0.0	0.0	1.6	2.0	2.1	3.2	0.0	0.0	0.0	5.3	3.7	0.0	0.0	3.0
AT4G34030.1	0.0	0.0	0.0	0.0	63.5	33.4	95.8	124.2	0.0	0.0	0.0	0.0	57.5	39.8	0.0	0.0
AT5G09650.1	0.0	0.0	0.0	0.0	42.3	15.7	19.2	22.1	0.0	0.0	0.0	0.0	0.0	0.0	3.8	6.9
AT1G48030.1	0.0	0.0	0.0	0.0	47.0	13.8	7.4	23.5	30.5	0.0	0.0	16.4	0.0	0.0	0.0	0.0
AT1G78370.1	0.0	0.0	0.0	0.0	23.5	27.5	9.6	24.8	0.0	0.0	16.7	24.6	0.0	0.0	0.0	0.0
AT1G48630.1	0.0	0.0	0.0	0.0	1.4	0.0	6.3	0.0	0.0	0.0	0.0	0.0	22.8	20.8	9.4	35.4
AT5G66570.1	0.0	0.0	0.0	1.9	30.6	36.4	6.4	4.8	15.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT4G24190.1	0.0	0.0	0.0	0.0	0.0	0.0	12.0	0.0	7.6	0.0	0.0	0.0	1.9	21.2	8.0	8.7
AT5G14320.1	22.8	14.5	154.3	33.4	0.0	0.0	3.2	2.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT2G37660.1	0.0	0.0	0.0	0.0	14.1	11.8	14.9	15.2	7.6	3.1	0.0	0.0	0.0	0.0	0.0	0.0
AT3G20820.1	0.0	0.0	0.0	0.0	0.0	0.0	12.8	12.4	0.0	0.0	0.0	8.2	0.0	6.6	8.9	4.6
AT4G24820.1	0.0	0.0	0.0	0.0	0.0	0.0	4.3	2.8	0.0	6.3	0.0	0.0	7.4	17.2	11.5	0.0
AT1G65960.2	0.0	0.0	0.0	0.0	2.4	0.0	5.3	4.1	0.0	0.0	0.0	0.0	5.6	15.9	12.8	0.0
AT4G15210.1	0.0	0.0	0.0	0.0	0.0	19.7	0.0	0.0	0.0	0.0	29.3	36.9	7.4	0.0	2.6	4.6
AT1G24510.1	0.0	0.0	0.0	0.0	7.1	2.0	4.3	0.0	0.0	0.0	0.0	0.0	9.3	17.2	6.4	0.0
AT2G16600.1	0.0	0.0	0.0	0.0	16.5	15.7	3.2	9.7	0.0	0.0	16.7	0.0	0.0	0.0	0.0	2.3
AT3G51260.1	0.0	0.0	0.0	0.0	16.5	2.0	5.3	4.8	0.0	0.0	0.0	0.0	11.1	9.3	0.0	0.0
AT2G35410.1	0.0	0.0	60.4	44.5	0.0	0.0	3.2	2.8	3.8	6.3	0.0	0.0	0.0	0.0	0.0	0.0
AT3G11940.1	0.0	0.0	0.0	0.0	2.4	0.0	6.4	0.0	0.0	12.6	25.1	26.2	0.0	2.7	0.0	0.0
AT4G33030.1	0.0	0.0	0.0	0.0	0.0	0.0	4.3	0.0	0.0	6.3	0.0	8.2	13.0	5.3	7.7	0.0
AT3G49120.1	0.0	0.0	0.0	0.0	10.6	19.7	1.6	2.8	0.0	0.0	0.0	0.0	11.1	0.0	1.3	0.0
AT1G75460.1	0.0	0.0	0.0	0.0	0.0	0.0	6.4	4.1	0.0	12.6	0.0	0.0	13.0	1.3	0.0	6.9
AT4G24830.1	0.0	0.0	26.8	44.5	4.7	0.0	1.1	1.4	0.0	0.0	0.0	0.0	0.0	0.0	5.1	0.0
AT1G07890.1	0.0	0.0	0.0	0.0	18.8	0.0	3.2	9.7	3.8	0.0	4.2	0.0	5.6	0.0	0.0	0.0
AT4G34670.1	0.0	0.0	0.0	0.0	0.0	0.0	9.8	4.6	5.0	0.0	0.0	0.0	5.6	3.7	3.2	0.0
AT5G19550.1	0.0	0.0	0.0	0.0	8.5	3.9	4.7	3.6	0.0	0.0	12.6	24.6	0.0	0.0	0.0	0.0
AT3G49010.1	0.0	0.0	0.0	0.0	0.0	0.0	10.6	2.8	0.0	0.0	0.0	0.0	5.6	4.0	1.3	2.3
AT3G52960.1	0.0	0.0	47.0	18.5	0.0	3.9	1.1	0.0	0.0	9.4	0.0	0.0	1.9	0.0	0.0	0.0
AT3G61440.1	0.0	0.0	0.0	0.0	9.4	5.9	4.3	1.4	19.0	0.0	0.0	0.0	0.0	0.0	2.6	0.0
AT5G48540.1	0.0	0.0	0.0	0.0	0.0	0.0	5.3	8.3	0.0	0.0	12.6	4.1	3.7	2.7	0.0	0.0
AT4G34450.1	0.0	0.0	0.0	0.0	0.0	0.0	6.4	0.0	0.0	3.1	8.4	0.0	1.9	5.3	5.1	0.0
AT5G54960.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	11.4	6.3	8.4	0.0	7.4	4.0	5.1	0.0
AT1G72730.1	0.0	0.0	0.0	0.0	0.0	0.0	4.5	1.8	0.0	4.1	4.2	0.0	4.3	10.2	0.0	0.0
AT5G10360.1	0.0	0.0	0.0	0.0	0.0	0.0	2.7	2.8	0.0	0.0	8.4	0.0	5.6	8.0	0.0	4.6
AT1G33590.1	0.0	0.0	0.0	0.0	2.4	5.9	2.1	9.7	7.6	0.0	0.0	8.2	0.0	0.0	0.0	0.0
AT2G30200.1	0.0	0.0	0.0	0.0	0.0	5.9	5.3	4.1	11.4	3.1	0.0	0.0	0.0	0.0	1.3	0.0
AT1G42960.1	0.0	0.0	0.0	0.0	0.0	0.0	1.1	0.0	0.0	0.0	0.0	12.3	9.3	4.0	3.8	2.3
AT4G09650.1	0.0	0.0	0.0	7.4	11.8	5.9	1.1	0.0	0.0	0.0	12.6	8.2	0.0	0.0	0.0	0.0
AT3G16640.1	0.0	0.0	0.0	0.0	9.4	7.9	4.3	2.8	0.0	0.0	0.0	4.1	0.0	0.0	1.3	0.0
AT3G25920.1	22.8	0.0	0.0	7.4	0.0	0.0	6.4	4.1	0.0	3.1	0.0	0.0	1.9	0.0	0.0	0.0
AT2G29630.1	0.0	0.0	0.0	0.0	0.0	7.9	1.1	2.8	0.0	0.0	0.0	0.0	0.0	2.7	2.6	9.1
AT5G59880.1	0.0	0.0	0.0	0.0	9.4	5.9	0.0	5.5	0.0	6.3	0.0	4.1	1.9	0.0	0.0	0.0
AT3G17820.1	0.0	0.0	0.0	0.0	7.1	2.0	0.0	1.4	9.1	15.7	0.0	0.0	0.0	0.0	2.6	0.0
AT1G54270.1	0.0	0.0	0.0	0.0	0.0	0.0	0.7	1.8	0.0	4.1	0.0	0.0	4.3	2.3	8.3	0.0
AT1G29880.1	0.0	0.0	0.0	0.0	0.0	9.8	3.2	0.0	7.6	3.1	4.2	0.0	0.0	0.0	0.0	2.3
AT5G65010.1	0.0	0.0	0.0	0.0	4.7	2.0	5.3	1.4	0.0	0.0	0.0	0.0	3.7	2.7	0.0	0.0
AT3G15020.1	0.0	0.0	0.0	0.0	1.6	0.0	0.0	1.4	0.0	8.8	10.5	11.9	0.0	0.0	3.6	0.0
AT4G19006.1	0.0	0.0	0.0	0.0	2.4	0.0	2.1	1.4	0.0	0.0	0.0	0.0	1.9	6.6	3.2	0.0
AT3G07100.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	6.3	4.2	0.0	3.7	1.3	5.1	4.6
AT4G38510.1	0.0	0.0	0.0	0.0	7.1	7.9	1.1	1.8	0.0	0.0	0.0	0.0	1.9	1.7	0.0	0.0
AT3G24503.1	0.0	0.0	0.0	0.0	7.1	0.0	1.1	2.8	0.0	0.0	0.0	4.1	3.7	0.0	1.3	0.0
AT3G54440.1	0.0	0.0	0.0	0.0	4.7	2.0	2.1	0.0	7.6	0.0	0.0	4.1	0.0	0.0	2.6	0.0
AT5G42790.1	0.0	0.0	0.0	0.0	3.5	0.0	0.0	0.0	3.8	0.0	6.3	0.0	2.8	2.0	3.2	0.0
AT5G12040.1	0.0	0.0	0.0	0.0	2.4	3.9	2.1	0.0	7.6	0.0	0.0	0.0	0.0	1.3	0.0	2.3
AT5G37830.1	0.0	0.0	0.0	0.0	0.0	2.0	1.1	2.8	0.0	0.0	0.0	4.1	0.0	0.0	2.6	2.3
AT1G56410.1	0.0	0.0	0.0	0.0	0.0	0.0	1.3	0.6	6.9	0.0	5.0	8.6	0.0	1.3	0.0	0.0
AT5G57950.1	0.0	0.0	6.7	0.0	0.0	0.0	0.0	1.4	3.8	0.0	4.2	0.0	3.7	0.0	0.0	0.0
AT5G19990.1	0.0	0.0	0.0	0.0	2.4	0.0	1.4	1.1	0.0	0.0	0.0	0.0	1.9	0.8	1.4	0.0
AT5G20000.1	0.0	0.0	0.0	0.0	2.4	0.0	1.4	1.1	0.0	0.0	0.0	0.0	1.9	0.8	1.4	0.0
AT5G44340.1	0.0	0.0	0.0	0.0	0.0	0.0	15.2	0.0	0.0	3.8	0.0	0.0	3.9	15.6	25.6	0.0
AT3G03250.1	0.0	0.0	0.0	0.0	40.0	23.0	6.4	15.2	0.0	0.0	0.0	0.0	0.0	0.0	2.6	0.0
AT1G52400.1	0.0	0.0	0.0	0.0	0.0	5.9	2.1	0.0	0.0	0.0	0.0	0.0	9.3	10.6	15.3	0.0
AT1G09640.1	0.0	0.0	0.0	0.0	10.8	0.0	5.7	0.0	0.0	0.0	0.0	0.0	11.9	9.9	7.7	0.0
AT3G08530.1	0.0	0.0	0.0	0.0	0.0	0.0	2.1	0.0	0.0	0.0	0.0	18.5	3.7	13.7	13.1	0.0
AT4G38740.1	0.0	0.0	0.0	0.0	16.5	13.8	1.1	15.2	0.0	0.0	0.0	4.1	0.0	0.0	0.0	0.0
AT5G51120.1	0.0	0.0	0.0	0.0	4.7	0.0	2.1	0.0	0.0	0.0	0.0	0.0	7.4	14.6	9.6	0.0
AT1G20010.1	0.0	0.0	0.0	0.0	0.0	0.0	5.4	0.0	0.0	5.3	0.0	0.0	5.2	19.4	2.6	0.0
AT3G20390.1	0.0	0.0	0.0	0.0	18.8	13.8	8.5	2.8	0.0	0.0	4.2	0.0	0.0	0.0	0.0	0.0
AT3G47290.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.4	0.0	31.4	25.1	32.8	0.0	0.0	1.3	0.0

Supplemental Table 2 (continued)

Accession (experiment # in PPDB: 2164 2165 2190 2191 2192 2193 2224 2225 2208 2209 2210 2211 2214 2215 2216 2217)	10000* NadjSP C exp. 2164 P3 control stroma	10000* NadjSP C exp. 2165 P3 S164A stroma	10000* NadjSP C exp. 2190 P3 control stroma	10000* NadjSP C exp. 2191 P3 S164A stroma	10000* NadjSP C exp. 2192 P3 S164A	10000* NadjSP C exp. 2193 P3 control	10000* NadjSP C exp. 2224 P3 control	10000* NadjSP C exp. 2225 P3 S164A	10000* NadjSP C Exp. 2208 P5 control	10000* NadjSP C exp. 2209 P5 S193A	10000* NadjSP C exp. 2210 P5 control	10000* NadjSP C exp. 2211 P5 S193A	10000* NadjSP C exp. 2214 P5 control	10000* NadjSP C exp. 2215 P5 S193A	10000* NadjSP C exp. 2216 P5 control	10000* NadjSP C exp. 2217 P5 S193A
AT4G10320.1	0.0	0.0	0.0	0.0	0.0	0.0	4.3	4.1	0.0	0.0	0.0	0.0	0.0	4.0	16.6	4.6
AT2G35370.1	0.0	0.0	0.0	0.0	12.2	9.8	6.9	8.3	0.0	0.0	0.0	0.0	0.0	0.0	2.6	0.0
AT4G37800.1	0.0	0.0	0.0	0.0	0.0	2.0	11.7	11.0	3.8	0.0	0.0	0.0	0.0	0.0	0.0	4.6
AT3G07390.1	0.0	0.0	0.0	0.0	4.7	23.6	1.1	5.5	0.0	12.6	0.0	0.0	0.0	0.0	0.0	0.0
AT3G18130.1	0.0	0.0	0.0	0.0	1.4	0.0	6.3	16.4	0.0	0.0	0.0	0.0	2.4	0.0	2.9	0.0
AT5G20890.1	0.0	0.0	0.0	0.0	0.0	0.0	5.3	2.8	0.0	0.0	0.0	0.0	3.7	13.3	2.6	0.0
AT4G31700.1	0.0	0.0	0.0	0.0	4.7	0.0	10.1	0.0	15.2	9.4	8.4	0.0	0.0	0.0	0.0	0.0
AT3G29360.1	0.0	0.0	0.0	0.0	0.0	3.9	4.8	0.0	0.0	0.0	0.0	0.0	0.0	3.3	9.6	6.9
AT3G53460.1	0.0	0.0	0.0	18.5	7.1	7.9	3.2	5.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT4G20890.1	0.0	0.0	0.0	0.0	0.0	0.0	4.9	0.0	0.0	3.8	0.0	0.0	3.9	11.1	3.2	0.0
AT3G13460.1	0.0	0.0	0.0	0.0	0.0	0.0	6.4	1.4	0.0	0.0	0.0	0.0	1.9	9.3	3.8	0.0
AT1G67280.1	0.0	0.0	0.0	0.0	11.1	8.5	3.2	3.2	11.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT5G28500.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	4.2	8.2	9.3	4.0	7.7	0.0
AT5G63310.1	0.0	0.0	0.0	0.0	7.1	5.9	5.3	6.9	0.0	0.0	0.0	4.1	0.0	0.0	0.0	0.0
AT5G47210.1	0.0	0.0	0.0	7.4	4.7	3.9	9.6	2.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT3G63190.1	0.0	0.0	0.0	11.1	2.4	9.8	2.1	6.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT1G20340.1	0.0	0.0	0.0	0.0	9.4	3.9	3.2	6.9	0.0	0.0	0.0	8.2	0.0	0.0	0.0	0.0
AT5G51820.1	0.0	0.0	0.0	0.0	9.4	7.9	2.1	5.5	3.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT1G04420.1	0.0	0.0	0.0	0.0	4.7	3.9	3.2	8.3	0.0	3.1	0.0	0.0	0.0	0.0	0.0	0.0
AT4G18440.1	0.0	0.0	0.0	0.0	2.4	7.9	0.0	5.5	0.0	0.0	0.0	0.0	0.0	2.7	3.8	0.0
AT4G08390.1	0.0	0.0	0.0	0.0	7.1	3.9	1.1	8.3	7.6	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT5G16715.1	0.0	0.0	0.0	0.0	0.0	2.0	3.2	0.0	0.0	0.0	0.0	0.0	0.0	2.7	3.8	9.1
AT1G35720.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	15.2	12.6	0.0	4.1	3.7	2.7	0.0	0.0
ATCG00770.1	0.0	0.0	6.7	11.1	0.0	0.0	3.2	2.8	0.0	9.4	0.0	0.0	0.0	0.0	0.0	0.0
AT2G27680.1	0.0	0.0	0.0	3.7	0.0	0.0	1.1	0.0	7.6	0.0	0.0	0.0	0.0	5.3	5.1	0.0
AT3G06650.1	0.0	0.0	0.0	0.0	4.7	5.9	0.0	0.0	5.7	0.0	0.0	0.0	4.6	4.0	0.0	0.0
AT2G41475.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.4	0.0	0.0	8.4	8.2	9.3	0.0	2.6	0.0
AT4G09040.1	0.0	0.0	20.1	18.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.9	1.3	0.0	2.3
AT5G41520.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.8	0.0	3.1	0.0	0.0	0.0	2.7	3.8	6.9
AT3G45030.1	0.0	0.0	0.0	0.0	3.5	2.9	3.2	4.1	0.0	0.0	0.0	6.2	0.0	0.0	0.0	0.0
AT5G62300.1	0.0	0.0	0.0	0.0	3.5	2.9	3.2	4.1	0.0	0.0	0.0	6.2	0.0	0.0	0.0	0.0
ATCG00160.1	0.0	0.0	13.4	14.8	0.0	0.0	2.1	1.4	0.0	0.0	0.0	0.0	0.0	0.0	1.3	0.0
ATCG00820.1	0.0	0.0	13.4	7.4	2.4	0.0	3.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.6	0.0
AT3G16470.1	0.0	0.0	0.0	0.0	0.0	3.9	0.0	4.1	0.0	0.0	0.0	4.1	3.7	2.7	0.0	0.0
AT2G36880.1	0.0	0.0	0.0	0.0	0.0	2.0	2.7	3.9	0.0	0.0	0.0	0.0	2.6	0.0	2.6	0.0
AT1G20260.1	0.0	0.0	0.0	0.0	0.0	0.0	1.1	1.8	0.0	0.0	0.0	0.0	5.6	1.7	3.8	0.0
AT1G76030.1	0.0	0.0	0.0	0.0	0.0	0.0	1.1	1.8	0.0	0.0	0.0	0.0	5.6	1.7	3.8	0.0
AT3G13120.1	0.0	0.0	6.7	7.4	0.0	0.0	0.0	1.4	0.0	0.0	12.6	0.0	0.0	0.0	2.6	0.0
AT1G80600.1	0.0	0.0	0.0	0.0	0.0	2.0	2.1	5.5	0.0	3.1	0.0	0.0	0.0	0.0	1.3	0.0
AT5G63890.1	0.0	0.0	0.0	0.0	4.7	7.9	0.0	0.0	3.8	0.0	4.2	4.1	0.0	0.0	0.0	0.0
AT2G01140.1	0.0	0.0	0.0	0.0	0.0	5.9	0.0	0.0	3.8	6.3	0.0	8.2	0.0	0.0	1.3	0.0
AT1G14610.1	0.0	0.0	0.0	0.0	0.0	2.0	3.2	1.4	0.0	0.0	0.0	0.0	0.0	2.7	2.6	0.0
AT5G62200.1	0.0	0.0	0.0	0.0	0.0	0.0	2.1	0.0	0.0	6.3	0.0	4.1	5.6	1.3	0.0	0.0
AT2G38540.1	0.0	0.0	0.0	0.0	4.7	5.9	1.1	2.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.3
AT3G12145.1	0.0	0.0	0.0	0.0	0.0	0.0	2.1	2.8	11.4	0.0	0.0	4.1	0.0	0.0	0.0	2.3
AT1G07770.1	0.0	0.0	0.0	0.0	4.7	0.0	2.7	1.8	0.0	0.0	0.0	0.0	2.8	0.0	1.3	0.0
AT5G59850.1	0.0	0.0	0.0	0.0	4.7	0.0	2.7	1.8	0.0	0.0	0.0	0.0	2.8	0.0	1.3	0.0
AT5G54600.1	0.0	0.0	6.7	3.7	0.0	0.0	2.1	2.8	0.0	0.0	0.0	0.0	0.0	0.0	2.6	0.0
AT2G19520.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.4	0.0	0.0	0.0	0.0	3.7	2.7	2.6	2.3
AT2G22230.1	0.0	0.0	13.4	7.4	2.4	0.0	0.0	0.0	3.8	0.0	0.0	4.1	0.0	0.0	0.0	0.0
AT4G35450.1	0.0	0.0	0.0	0.0	4.7	0.0	1.1	2.8	0.0	0.0	0.0	0.0	0.0	1.3	1.3	0.0
AT5G10770.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3.1	4.2	0.0	3.7	2.7	1.3	0.0
AT3G49110.1	0.0	0.0	0.0	0.0	1.2	2.0	1.6	2.8	0.0	0.0	0.0	0.0	0.0	0.0	1.3	0.0
AT5G50250.1	0.0	0.0	0.0	0.0	0.0	0.0	1.1	1.4	0.0	0.0	0.0	0.0	1.9	1.3	0.0	2.3
AT1G09780.1	0.0	0.0	0.0	0.0	34.6	29.5	9.4	24.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT1G43170.1	0.0	0.0	0.0	0.0	0.0	0.0	17.0	0.0	0.0	44.0	0.0	0.0	0.0	10.6	20.4	0.0
AT5G55070.1	0.0	0.0	0.0	0.0	0.0	0.0	17.1	5.5	0.0	0.0	0.0	0.0	0.0	10.6	14.0	0.0
AT1G45000.1	0.0	0.0	0.0	0.0	0.0	0.0	3.5	0.0	0.0	0.0	0.0	0.0	2.8	27.6	15.7	0.0
AT1G78380.1	0.0	0.0	0.0	0.0	16.5	17.7	7.4	15.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT4G02930.1	0.0	0.0	0.0	0.0	0.0	0.0	11.7	0.0	0.0	0.0	0.0	0.0	3.7	15.9	11.5	0.0
AT3G02530.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.1	0.0	0.0	0.0	0.0	20.4	15.9	11.5	0.0
AT5G62690.1	0.0	0.0	0.0	0.0	0.0	0.0	6.8	0.0	0.0	3.8	0.0	0.0	0.0	28.8	2.6	0.0
AT5G62700.1	0.0	0.0	0.0	0.0	0.0	0.0	6.8	0.0	0.0	3.8	0.0	0.0	0.0	28.8	2.6	0.0
AT5G02870.1	0.0	0.0	0.0	0.0	0.0	0.0	10.6	0.0	0.0	15.7	0.0	0.0	0.0	8.6	10.8	0.0
AT2G38040.1	0.0	0.0	0.0	74.2	0.0	0.0	5.3	2.8	0.0	0.0	0.0	0.0	0.0	2.7	0.0	0.0
AT1G16080.1	0.0	0.0	0.0	0.0	23.5	15.7	4.3	8.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT1G70580.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	11.4	47.2	18.8	16.4	0.0	0.0	0.0	0.0
AT4G34870.1	0.0	0.0	0.0	0.0	18.8	21.6	2.1	4.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT1G32470.1	0.0	0.0	0.0	0.0	18.3	9.8	5.9	4.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
ATCG00170.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	11.4	9.4	0.0	0.0	0.0	6.6	11.5	0.0
AT1G29900.1	0.0	0.0	0.0	0.0	0.0	0.0	6.4	0.0	7.6	0.0	0.0	12.3	0.0	0.0	11.5	0.0
AT5G09900.1	0.0	0.0	0.0	0.0	0.0	0.0	4.3	0.0	0.0	0.0	0.0	0.0	1.9	8.9	10.2	0.0
AT5G16050.1	0.0	0.0	0.0	0.0	0.0	0.0	8.1	7.2	15.2	0.0	0.0	0.0	2.6	0.0	0.0	0.0
AT3G48420.1	0.0	0.0	0.0	0.0	16.5	9.8	2.1	5.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT1G53750.1	0.0	0.0	0.0	0.0	0.0	0.0	5.3	0.0	0.0	0.0	0.0	0.0	3.7	9.3	5.1	0.0
AT2G43710.1	0.0	0.0	0.0	0.0	2.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	9.3	6.6	7.7	0.0
AT1G04810.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.8	5.7	0.0	0.0	0.0	4.6	0.0	13.9	0.0
AT3G26060.1	0.0	0.0	20.1	14.8	0.0	0.0	5.3	5.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Supplemental Table 2 (continued)

Accession (experiment # in PPDB: 2164 2165 2190 2191 2192 2193 2224 2225 2208 2209 2210 2211 2214 2215 2216 2217)	10000* NadjSP C exp. 2164 P3 control stroma	10000* NadjSP C exp. 2165 P3 S164A stroma	10000* NadjSP C exp. 2190 P3 control stroma	10000* NadjSP C exp. 2191 P3 S164A stroma	10000* NadjSP C exp. 2192 P3 S164A	10000* NadjSP C exp. 2193 P3 control	10000* NadjSP C exp. 2224 P3 control	10000* NadjSP C exp. 2225 P3 S164A	10000* NadjSP C Exp. 2208 P5 control	10000* NadjSP C exp. 2209 P5 S193A	10000* NadjSP C exp. 2210 P5 control	10000* NadjSP C exp. 2211 P5 S193A	10000* NadjSP C exp. 2214 P5 control	10000* NadjSP C exp. 2215 P5 S193A	10000* NadjSP C exp. 2216 P5 control	10000* NadjSP C exp. 2217 P5 S193A
AT5G46290.1	0.0	0.0	0.0	11.1	0.0	0.0	6.4	0.0	0.0	0.0	0.0	0.0	0.0	2.7	6.4	0.0
AT1G50480.1	0.0	0.0	0.0	0.0	0.0	7.9	7.4	5.5	0.0	3.1	0.0	0.0	0.0	0.0	0.0	0.0
AT5G14910.1	0.0	0.0	26.8	18.5	0.0	3.9	4.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT1G74260.1	0.0	0.0	0.0	0.0	0.0	0.0	5.3	6.9	11.4	0.0	0.0	0.0	0.0	0.0	2.6	0.0
AT1G80410.1	0.0	0.0	0.0	0.0	0.0	0.0	5.3	2.8	0.0	0.0	0.0	0.0	0.0	5.3	5.1	0.0
AT5G24770.1	0.0	0.0	0.0	0.0	0.0	2.6	0.0	3.6	3.8	0.0	0.0	0.0	15.2	0.0	0.0	0.0
AT5G16710.1	0.0	0.0	0.0	3.7	4.7	13.8	0.0	4.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT3G08030.1	0.0	0.0	0.0	0.0	2.4	5.9	4.3	5.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT3G59970.3	0.0	0.0	0.0	0.0	7.1	11.8	1.1	2.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT3G01340.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.8	0.0	0.0	0.0	0.0	5.6	0.0	6.4	4.6
AT2G37270.1	0.0	0.0	0.0	0.0	2.4	0.0	4.3	0.0	0.0	0.0	0.0	6.6	0.0	0.0	5.1	0.0
AT1G50200.1	0.0	0.0	0.0	0.0	0.0	0.0	1.1	0.0	0.0	0.0	0.0	0.0	3.7	0.0	3.8	9.1
AT4G09010.1	0.0	0.0	0.0	0.0	0.0	3.9	4.3	0.0	0.0	0.0	0.0	0.0	3.7	2.7	0.0	0.0
AT3G46970.1	0.0	0.0	0.0	0.0	0.0	0.0	3.2	0.0	15.2	3.1	0.0	0.0	0.0	0.0	2.6	0.0
AT5G23540.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	6.3	0.0	0.0	0.0	5.3	3.8	2.3
AT4G29010.1	0.0	0.0	0.0	0.0	0.0	0.0	3.2	5.5	0.0	6.3	0.0	0.0	0.0	0.0	0.0	2.3
AT4G26690.1	0.0	0.0	0.0	0.0	4.7	4.9	2.1	4.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT3G16400.1	0.0	0.0	0.0	0.0	0.0	7.9	0.0	1.4	0.0	0.0	0.0	0.0	3.7	0.0	3.8	0.0
AT1G33120.1	0.0	0.0	0.0	0.0	0.0	0.0	4.8	4.1	0.0	0.0	0.0	0.0	1.9	0.0	0.0	2.3
AT1G33140.1	0.0	0.0	0.0	0.0	0.0	0.0	4.8	4.1	0.0	0.0	0.0	0.0	1.9	0.0	0.0	2.3
AT3G02080.1	0.0	0.0	0.0	0.0	0.0	0.0	1.1	0.0	0.0	0.0	4.2	16.4	0.0	0.0	4.5	0.0
AT4G19410.2	0.0	0.0	0.0	0.0	1.2	2.0	7.4	0.0	0.0	0.0	4.2	0.0	0.0	0.0	0.0	0.0
AT1G48350.1	34.2	0.0	13.4	7.4	0.0	0.0	0.0	2.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT5G13650.1	0.0	0.0	0.0	0.0	0.0	0.0	2.1	1.4	0.0	0.0	0.0	0.0	0.0	4.0	3.8	0.0
AT1G20200.1	0.0	0.0	0.0	0.0	0.0	0.0	2.1	1.4	0.0	0.0	0.0	0.0	0.0	1.3	6.4	0.0
AT2G47470.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	4.1	3.7	0.0	1.3	11.4
AT1G65900.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.4	0.0	0.0	0.0	0.0	7.4	4.0	1.3	0.0
AT5G19240.1	0.0	0.0	0.0	0.0	0.0	0.0	2.1	4.1	0.0	0.0	0.0	0.0	3.7	0.0	2.6	0.0
AT1G03630.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3.7	2.7	2.6	4.6
AT4G03520.1	0.0	0.0	0.0	0.0	0.0	0.0	2.1	2.8	0.0	0.0	0.0	12.3	1.9	0.0	0.0	0.0
AT4G39330.1	0.0	0.0	0.0	0.0	0.0	5.9	3.2	1.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.3
AT3G18060.1	0.0	0.0	0.0	0.0	0.0	9.8	1.1	1.4	0.0	0.0	0.0	0.0	0.0	0.0	1.3	0.0
AT1G26630.1	0.0	0.0	0.0	0.0	1.6	2.0	4.3	2.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT1G68590.1	0.0	0.0	6.7	7.4	0.0	0.0	1.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3.8	0.0
AT4G34620.1	0.0	0.0	13.4	0.0	0.0	0.0	0.0	4.1	0.0	0.0	0.0	4.1	0.0	0.0	0.0	2.3
AT1G62640.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3.7	2.7	2.6	2.3
AT1G20160.1	0.0	0.0	0.0	0.0	0.0	0.0	1.1	2.8	0.0	0.0	0.0	0.0	3.7	0.0	2.6	0.0
AT2G44350.1	0.0	0.0	0.0	0.0	4.7	3.9	2.1	0.0	0.0	3.1	0.0	0.0	0.0	0.0	0.0	0.0
AT2G41530.1	0.0	0.0	0.0	0.0	4.7	3.9	2.1	0.0	3.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT2G21250.1	0.0	0.0	0.0	0.0	2.4	3.9	0.0	2.8	5.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT1G76080.1	0.0	0.0	0.0	0.0	4.7	2.0	1.1	2.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT1G19670.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	4.2	0.0	3.7	2.7	1.3	0.0
AT3G52140.1	0.0	0.0	0.0	0.0	0.0	0.0	2.1	0.0	3.8	0.0	0.0	0.0	0.0	1.3	2.6	0.0
AT3G28220.1	0.0	0.0	0.0	0.0	0.0	3.9	1.1	1.4	0.0	0.0	0.0	0.0	1.9	0.0	0.0	0.0
AT3G08740.1	0.0	0.0	6.7	7.4	0.0	2.0	0.0	1.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT2G42740.1	0.0	0.0	0.0	0.0	0.7	0.6	1.4	0.0	0.0	0.0	0.0	0.0	1.5	0.0	0.0	0.0
AT3G58700.1	0.0	0.0	0.0	0.0	0.7	0.6	1.4	0.0	0.0	0.0	0.0	0.0	1.5	0.0	0.0	0.0
AT4G18730.1	0.0	0.0	0.0	0.0	0.7	0.6	1.4	0.0	0.0	0.0	0.0	0.0	1.5	0.0	0.0	0.0
AT5G45775.1	0.0	0.0	0.0	0.0	0.7	0.6	1.4	0.0	0.0	0.0	0.0	0.0	1.5	0.0	0.0	0.0
AT5G13630.1	0.0	0.0	0.0	0.0	0.0	0.0	8.5	0.0	0.0	0.0	0.0	0.0	0.0	1.3	28.1	0.0
AT1G70730.1	0.0	0.0	0.0	0.0	0.0	13.4	7.0	12.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT1G76160.1	0.0	0.0	0.0	0.0	14.1	25.6	0.0	2.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT3G08590.1	0.0	0.0	0.0	0.0	24.2	17.7	1.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT2G05990.1	0.0	0.0	0.0	0.0	0.0	0.0	6.4	5.5	30.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT1G80380.2	0.0	0.0	0.0	0.0	0.0	0.0	4.3	0.0	0.0	0.0	0.0	0.0	0.0	11.9	3.8	0.0
AT3G58500.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.9	13.3	6.4	0.0
AT5G12250.1	0.0	0.0	0.0	0.0	0.0	0.0	2.7	0.0	0.0	5.3	0.0	0.0	0.0	0.0	14.5	0.0
AT3G23940.2	0.0	0.0	0.0	0.0	9.4	15.7	0.0	4.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT5G05780.1	0.0	0.0	0.0	0.0	0.0	0.0	5.3	0.0	0.0	0.0	0.0	0.0	0.0	10.6	0.0	4.6
AT1G18540.1	0.0	0.0	0.0	0.0	0.0	0.0	8.0	0.0	15.2	0.0	0.0	0.0	4.6	0.0	0.0	0.0
AT4G19210.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	4.0	11.5	4.6
AT5G03340.1	0.0	0.0	0.0	0.0	0.0	0.0	6.4	5.1	0.0	0.0	16.7	0.0	0.0	0.0	0.0	0.0
AT5G04740.1	0.0	0.0	13.4	37.1	0.0	0.0	0.0	1.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT2G30870.1	0.0	0.0	0.0	0.0	0.0	7.9	4.3	6.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT4G11010.1	0.0	0.0	0.0	0.0	14.1	11.8	1.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT5G28840.1	0.0	0.0	0.0	0.0	0.0	0.0	6.4	4.1	0.0	0.0	0.0	0.0	0.0	0.0	5.1	0.0
AT2G04842.1	0.0	0.0	0.0	0.0	0.0	0.0	4.3	0.0	0.0	0.0	0.0	0.0	0.0	5.3	5.1	0.0
AT5G52520.1	0.0	0.0	0.0	29.7	4.7	0.0	2.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT2G22360.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	6.3	0.0	0.0	0.0	5.3	7.7	0.0
AT5G64370.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	9.3	5.3	3.8	0.0
AT3G05970.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	11.9	2.6	2.3
AT4G25130.1	0.0	0.0	0.0	0.0	0.0	0.0	2.1	11.0	0.0	0.0	0.0	0.0	1.9	0.0	0.0	0.0
AT1G64520.1	0.0	0.0	0.0	0.0	0.0	0.0	2.1	0.0	0.0	0.0	4.2	0.0	0.0	10.6	0.0	0.0
AT2G19730.1	0.0	0.0	0.0	0.0	0.0	0.0	4.3	8.3	0.0	3.1	0.0	0.0	0.0	0.0	0.0	0.0
AT3G58140.1	0.0	0.0	0.0	0.0	0.0	0.0	2.1	0.0	0.0	0.0	0.0	0.0	0.0	1.3	8.9	0.0
AT4G04640.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3.8	0.0	0.0	0.0	0.0	9.3	2.6	0.0
AT5G60600.1	0.0	0.0	0.0	0.0	0.0	0.0	6.4	0.0	7.6	0.0	8.4	0.0	0.0	0.0	0.0	0.0
AT3G22110.1	0.0	0.0	0.0	0.0	4.7	0.0	3.2	6.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Supplemental Table 2 (continued)

Accession (experiment # in PPDB: 2164 2165 2190 2191 2192 2193 2224 2225 2208 2209 2210 2211 2214 2215 2216 2217)	10000* NadjSP C exp. 2164 P3 control stroma	10000* NadjSP C exp. 2165 P3 S164A stroma	10000* NadjSP C exp. 2190 P3 control stroma	10000* NadjSP C exp. 2191 P3 S164A stroma	10000* NadjSP C exp. 2192 P3 S164A	10000* NadjSP C exp. 2193 P3 control	10000* NadjSP C exp. 2224 P3 control	10000* NadjSP C exp. 2225 P3 S164A	10000* NadjSP C Exp. 2208 P5 control	10000* NadjSP C exp. 2209 P5 S193A	10000* NadjSP C exp. 2210 P5 control	10000* NadjSP C exp. 2211 P5 S193A	10000* NadjSP C exp. 2214 P5 control	10000* NadjSP C exp. 2215 P5 S193A	10000* NadjSP C exp. 2216 P5 control	10000* NadjSP C exp. 2217 P5 S193A
AT4G33510.1	0.0	0.0	0.0	11.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.6	6.8	0.0
AT1G21750.1	0.0	0.0	0.0	0.0	3.5	11.4	0.0	2.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT1G79790.1	0.0	0.0	0.0	0.0	0.0	0.0	2.1	5.5	0.0	0.0	0.0	0.0	5.6	0.0	0.0	0.0
AT3G29320.1	0.0	0.0	0.0	0.0	0.0	2.0	6.4	0.0	7.6	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT5G55220.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.4	0.0	0.0	0.0	0.0	1.9	0.0	8.9	0.0
AT5G61410.1	0.0	0.0	0.0	0.0	0.0	5.9	3.2	4.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT3G09630.1	0.0	0.0	0.0	0.0	0.0	0.0	5.3	0.0	0.0	0.0	0.0	0.0	0.0	3.3	1.9	0.0
AT5G62530.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	4.1	0.0	0.0	0.0	12.3	5.6	0.0	0.0	0.0
AT1G75380.1	0.0	0.0	0.0	0.0	0.0	0.0	5.3	0.0	0.0	0.0	0.0	0.0	0.0	2.7	2.6	0.0
AT3G05900.1	0.0	0.0	0.0	11.1	0.0	3.9	0.0	5.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT3G51330.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	8.4	0.0	11.1	1.3	0.0	0.0
AT4G31840.1	0.0	0.0	0.0	0.0	0.0	5.9	2.1	5.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT5G45620.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3.1	0.0	0.0	0.0	6.6	3.2	0.0
AT5G24490.1	0.0	0.0	0.0	0.0	0.0	0.0	5.3	0.0	0.0	0.0	0.0	0.0	1.9	0.0	2.6	0.0
ATCG00380.1	0.0	0.0	0.0	0.0	0.0	0.0	4.3	4.1	0.0	0.0	0.0	0.0	1.9	0.0	0.0	0.0
AT1G15140.1	0.0	0.0	0.0	0.0	0.0	2.0	1.1	8.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT1G02920.1	0.0	0.0	0.0	0.0	0.0	0.0	4.3	2.8	0.0	0.0	0.0	0.0	0.0	2.7	0.0	0.0
AT3G19710.1	0.0	0.0	0.0	0.0	0.0	3.9	0.0	0.0	0.0	0.0	0.0	16.4	0.0	2.7	0.0	0.0
AT3G60820.1	0.0	0.0	0.0	0.0	2.4	0.0	0.0	5.5	0.0	0.0	0.0	0.0	5.6	0.0	0.0	0.0
AT5G18380.1	0.0	0.0	0.0	0.0	0.0	0.0	2.1	4.1	0.0	0.0	0.0	12.3	0.0	0.0	0.0	0.0
AT1G24020.1	0.0	0.0	0.0	0.0	7.1	7.9	0.0	0.0	0.0	3.1	0.0	0.0	0.0	0.0	0.0	0.0
AT2G25060.1	0.0	0.0	0.0	0.0	7.1	5.9	0.0	0.0	7.6	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT2G35920.1	0.0	0.0	0.0	0.0	0.0	0.0	3.2	0.0	0.0	0.0	0.0	0.0	0.0	5.3	1.3	0.0
AT2G44120.1	0.0	0.0	0.0	0.0	0.0	0.0	2.9	0.0	0.0	0.0	0.0	0.0	6.5	0.0	1.3	0.0
ATCG00190.1	0.0	0.0	0.0	0.0	0.0	0.0	1.1	0.0	0.0	0.0	0.0	0.0	0.0	2.7	5.1	0.0
AT1G66430.1	0.0	0.0	0.0	0.0	0.0	3.9	4.3	0.0	3.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT5G49460.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	5.7	0.0	0.0	0.0	4.6	4.0	0.0	0.0
AT4G20260.1	0.0	0.0	0.0	0.0	7.1	0.0	1.1	4.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT1G07140.1	0.0	0.0	0.0	0.0	0.0	0.0	2.1	1.4	15.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT2G04160.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	12.6	0.0	1.9	0.0	3.8	0.0
AT1G04270.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.1	0.0	0.0	0.0	0.0	3.7	0.0	3.8	0.0
AT5G09510.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.1	0.0	0.0	0.0	0.0	3.7	0.0	3.8	0.0
AT1G63000.1	0.0	0.0	0.0	0.0	0.7	0.0	4.3	0.0	7.6	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT1G75330.1	0.0	0.0	0.0	0.0	0.0	0.0	1.1	4.1	7.6	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT4G23100.1	0.0	0.0	0.0	0.0	0.0	3.9	2.1	0.0	0.0	0.0	0.0	0.0	0.0	2.7	0.0	0.0
AT4G26300.1	0.0	0.0	0.0	14.8	0.0	2.0	0.0	1.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT3G48110.1	0.0	0.0	0.0	0.0	0.0	0.0	2.1	0.0	11.4	0.0	0.0	0.0	0.0	0.0	1.3	0.0
AT5G23120.1	0.0	0.0	0.0	0.0	7.1	0.0	2.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.3	0.0
AT4G27090.1	0.0	0.0	0.0	0.0	0.0	0.0	3.2	0.0	0.0	0.0	0.0	8.2	1.9	0.0	0.0	0.0
AT3G10920.1	0.0	0.0	0.0	0.0	2.4	0.0	1.1	5.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT2G29450.1	0.0	0.0	0.0	0.0	2.4	2.0	0.0	5.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT3G12290.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.8	7.6	0.0	0.0	8.2	0.0	0.0	0.0	0.0
AT1G78300.1	0.0	0.0	0.0	0.0	4.0	0.0	0.0	0.0	7.6	0.0	0.0	0.0	0.0	3.0	0.0	0.0
AT4G11150.1	0.0	0.0	0.0	0.0	0.0	0.0	3.2	2.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.3
AT2G26740.1	0.0	0.0	0.0	0.0	4.7	0.0	0.0	1.4	11.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT3G07720.1	0.0	0.0	0.0	0.0	4.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	4.0	1.3	0.0
AT5G15490.1	0.0	0.0	0.0	0.0	0.0	3.9	1.6	0.0	0.0	0.0	0.0	0.0	0.0	3.3	0.0	0.0
AT5G28050.1	0.0	0.0	0.0	0.0	0.0	0.0	2.1	2.8	0.0	0.0	0.0	0.0	3.7	0.0	0.0	0.0
AT5G60370.1	0.0	0.0	0.0	0.0	0.0	0.0	2.1	2.8	0.0	0.0	0.0	0.0	0.0	2.7	0.0	0.0
AT2G37620.1	0.0	0.0	0.0	0.0	0.0	0.0	2.3	0.0	0.0	0.0	0.0	0.0	0.0	2.9	1.7	0.0
AT3G53750.1	0.0	0.0	0.0	0.0	0.0	0.0	2.3	0.0	0.0	0.0	0.0	0.0	0.0	2.9	1.7	0.0
AT3G11510.1	0.0	0.0	0.0	0.0	0.7	0.0	4.3	1.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT3G52580.1	0.0	0.0	0.0	0.0	0.7	0.0	4.3	1.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT5G47190.1	0.0	14.5	0.0	3.7	0.0	0.0	3.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT2G04700.1	0.0	0.0	0.0	0.0	2.4	0.0	2.1	2.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT4G28706.1	0.0	0.0	0.0	3.7	0.0	0.0	2.1	2.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT4G35630.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.4	0.0	0.0	0.0	4.1	0.0	0.0	3.8	0.0
AT2G20260.1	0.0	0.0	0.0	0.0	4.7	3.9	0.0	1.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT1G17100.1	0.0	0.0	0.0	0.0	0.0	0.0	2.1	0.0	0.0	0.0	0.0	0.0	3.7	1.3	0.0	0.0
AT1G27450.1	0.0	0.0	0.0	0.0	2.4	3.9	0.0	2.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT3G48000.1	0.0	0.0	0.0	0.0	0.0	0.0	2.1	0.0	3.8	0.0	0.0	0.0	0.0	2.7	0.0	0.0
AT5G08300.1	0.0	0.0	0.0	0.0	7.1	1.0	0.0	2.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT3G52990.1	0.0	0.0	0.0	0.0	0.0	0.0	1.1	0.0	0.0	0.0	0.0	8.2	0.0	2.7	0.0	0.0
AT4G05420.1	0.0	0.0	0.0	0.0	0.0	0.0	2.1	0.0	0.0	0.0	0.0	0.0	1.9	2.7	0.0	0.0
AT5G07030.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.8	0.0	0.0	0.0	4.1	3.7	0.0	0.0	0.0
AT5G26710.1	0.0	0.0	0.0	0.0	0.0	2.0	3.2	0.0	3.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT1G02500.1	0.0	0.0	0.0	0.0	0.0	2.0	1.6	0.0	0.0	0.0	0.0	0.0	0.0	2.7	0.0	0.0
AT5G08670.1	0.0	0.0	0.0	0.0	1.6	4.9	0.0	0.0	0.0	0.0	0.0	4.5	0.0	0.0	0.0	0.0
AT5G08680.1	0.0	0.0	0.0	0.0	1.6	4.9	0.0	0.0	0.0	0.0	0.0	4.5	0.0	0.0	0.0	0.0
AT5G08690.1	0.0	0.0	0.0	0.0	1.6	4.9	0.0	0.0	0.0	0.0	0.0	4.5	0.0	0.0	0.0	0.0
AT3G16530.1	0.0	0.0	0.0	0.0	0.0	2.4	0.0	1.4	0.0	0.0	0.0	0.0	3.7	0.0	0.0	0.0
AT1G74050.1	0.0	0.0	0.0	0.0	0.0	0.0	1.4	0.0	0.0	0.0	0.0	6.2	2.4	0.0	0.0	0.0
AT1G74060.1	0.0	0.0	0.0	0.0	0.0	0.0	1.4	0.0	0.0	0.0	0.0	6.2	2.4	0.0	0.0	0.0
AT5G16660.1	11.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	8.2	1.9	0.0	0.0	0.0
AT1G64510.1	0.0	0.0	0.0	3.7	0.0	0.0	2.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.3	0.0
ATCG00420.1	0.0	0.0	0.0	0.0	0.0	0.0	1.1	1.4	0.0	0.0	0.0	0.0	0.0	0.0	2.6	0.0
AT2G27530.1	0.0	0.0	0.0	0.0	0.0	0.0	1.1	2.8	0.0	0.0	4.2	0.0	0.0	0.0	0.0	0.0

Supplemental Table 2 (continued)

Accession (experiment # in PPDB: 2164 2165 2190 2191 2192 2193 2224 2225 2208 2209 2210 2211 2214 2215 2216 2217)	10000* NadjSP C exp. 2164 P3 control stroma	10000* NadjSP C exp. 2165 P3 S164A stroma	10000* NadjSP C exp. 2190 P3 control stroma	10000* NadjSP C exp. 2191 P3 S164A stroma	10000* NadjSP C exp. 2192 P3 S164A	10000* NadjSP C exp. 2193 P3 control	10000* NadjSP C exp. 2224 P3 control	10000* NadjSP C exp. 2225 P3 S164A	10000* NadjSP C Exp. 2208 P5 control	10000* NadjSP C exp. 2209 P5 S193A	10000* NadjSP C exp. 2210 P5 control	10000* NadjSP C exp. 2211 P5 S193A	10000* NadjSP C exp. 2214 P5 control	10000* NadjSP C exp. 2215 P5 S193A	10000* NadjSP C exp. 2216 P5 control	10000* NadjSP C exp. 2217 P5 S193A
AT1G17290.1	0.0	0.0	0.0	0.0	2.4	0.0	0.0	0.0	0.0	3.1	8.4	0.0	0.0	0.0	0.0	0.0
AT1G30690.1	0.0	0.0	0.0	0.0	0.0	0.0	2.1	1.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.3
AT1G01900.1	0.0	0.0	0.0	0.0	0.0	0.0	1.1	1.4	0.0	0.0	0.0	0.0	3.7	0.0	0.0	0.0
AT1G48660.1	0.0	0.0	6.7	3.7	4.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT3G27670.1	0.0	0.0	0.0	0.0	0.0	0.0	1.1	0.0	0.0	0.0	0.0	0.0	0.0	2.7	1.3	0.0
AT3G56070.1	0.0	0.0	0.0	0.0	2.4	2.0	0.0	2.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT5G20830.1	0.0	0.0	0.0	0.0	0.0	0.0	1.1	1.4	0.0	0.0	0.0	0.0	0.0	2.7	0.0	0.0
AT5G64130.1	0.0	0.0	0.0	0.0	0.0	2.0	0.0	1.4	0.0	0.0	0.0	0.0	3.7	0.0	0.0	0.0
AT3G50820.1	0.0	0.0	0.0	1.9	0.0	2.9	0.0	2.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT1G21720.1	0.0	0.0	0.0	0.0	2.4	1.0	0.0	2.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT1G77440.1	0.0	0.0	0.0	0.0	2.4	1.0	0.0	2.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT5G65430.1	0.0	0.0	0.0	0.0	0.0	0.0	1.2	0.0	0.0	0.0	0.0	0.0	2.2	1.5	0.0	0.0
AT1G31230.1	0.0	0.0	0.0	3.7	0.0	0.0	0.0	0.0	0.0	3.1	0.0	0.0	0.0	0.0	1.3	0.0
AT2G25080.1	0.0	0.0	0.0	3.7	2.4	0.0	1.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT4G32520.1	0.0	0.0	6.7	3.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.3
AT4G31300.1	0.0	0.0	0.0	0.0	0.0	2.0	1.1	1.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT3G13860.1	0.0	0.0	0.0	0.0	0.0	2.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.3	1.3	0.0
AT1G64790.1	0.0	0.0	0.0	0.0	0.0	0.0	1.1	0.0	0.0	0.0	0.0	0.0	0.0	1.3	1.3	0.0
AT2G23600.1	0.0	0.0	0.0	0.0	2.4	0.0	0.0	0.0	0.0	0.0	0.0	4.1	1.9	0.0	0.0	0.0
AT3G12800.1	0.0	0.0	0.0	0.0	0.0	0.0	1.1	1.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.3
AT2G32240.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3.8	0.0	0.0	0.0	0.0	1.3	1.3	0.0
AT1G31180.1	0.0	0.0	0.0	0.0	0.0	2.0	0.0	0.0	0.0	3.1	0.0	0.0	0.0	0.0	0.9	0.0
AT1G36240.1	0.0	0.0	0.0	0.0	0.7	0.0	1.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.3	0.0
AT1G77940.1	0.0	0.0	0.0	0.0	0.7	0.0	1.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.3	0.0
AT3G18740.1	0.0	0.0	0.0	0.0	0.7	0.0	1.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.3	0.0
AT2G47170.1	0.0	0.0	0.0	0.0	1.2	1.0	1.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT1G10630.1	0.0	0.0	0.0	0.0	1.2	1.0	1.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT1G23490.1	0.0	0.0	0.0	0.0	1.2	1.0	1.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT1G70490.1	0.0	0.0	0.0	0.0	1.2	1.0	1.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT3G62290.1	0.0	0.0	0.0	0.0	1.2	1.0	1.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT5G14670.1	0.0	0.0	0.0	0.0	1.2	1.0	1.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT5G28150.1	0.0	0.0	0.0	3.7	1.2	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT1G36180.1	0.0	0.0	0.0	0.0	3.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	43.0	0.0
AT1G69830.1	0.0	0.0	0.0	89.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.3	0.0
AT1G16720.1	11.4	0.0	0.0	85.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT3G05530.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	13.3	8.9	0.0
AT2G21390.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.7	16.7	0.0
AT2G47610.1	0.0	0.0	0.0	0.0	0.0	0.0	14.9	0.0	0.0	3.1	0.0	0.0	0.0	0.0	0.0	0.0
AT5G15200.1	0.0	0.0	0.0	0.0	0.0	0.0	10.6	6.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT1G58370.1	0.0	0.0	0.0	0.0	0.0	0.0	11.7	4.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT4G34200.1	0.0	0.0	0.0	45.6	0.0	0.0	0.0	0.0	3.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT3G60770.1	0.0	0.0	0.0	0.0	0.0	0.0	11.7	0.0	0.0	0.0	0.0	4.1	0.0	0.0	0.0	0.0
AT1G34430.1	0.0	0.0	26.8	28.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT5G17310.2	0.0	0.0	0.0	0.0	0.0	18.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.6	0.0
AT4G24620.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	11.0	7.6	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT1G09620.1	0.0	0.0	0.0	0.0	0.0	0.0	3.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	8.9	0.0
AT1G79930.1	0.0	0.0	0.0	0.0	0.0	0.0	8.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.6	0.0
AT5G23820.1	0.0	0.0	0.0	0.0	4.7	15.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT5G35790.1	0.0	0.0	0.0	29.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.3	0.0	0.0
AT2G36460.1	0.0	0.0	0.0	0.0	0.0	0.0	5.0	0.0	0.0	0.0	0.0	0.0	0.0	5.7	0.0	0.0
AT2G02100.1	0.0	0.0	0.0	0.0	0.0	0.0	4.3	6.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT5G60360.1	0.0	0.0	0.0	0.0	0.0	0.0	3.2	8.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT5G27410.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	5.3	6.4	0.0
AT1G03220.1	0.0	0.0	0.0	0.0	7.1	9.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT1G12240.1	0.0	0.0	0.0	0.0	2.4	13.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT3G02870.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.4	26.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT1G29250.1	0.0	0.0	0.0	0.0	0.0	0.0	6.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.6	0.0
AT4G39980.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	6.4	3.4	0.0
AT3G03710.1	0.0	0.0	0.0	22.3	0.0	0.0	1.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT1G56450.1	0.0	0.0	0.0	0.0	0.0	0.0	2.1	6.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT3G49910.1	0.0	0.0	0.0	0.0	0.0	0.0	6.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.3	0.0
AT4G29350.1	0.0	0.0	0.0	0.0	7.1	7.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT5G14780.1	0.0	0.0	0.0	0.0	0.0	0.0	4.3	4.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT1G21880.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	15.2	0.0	0.0	0.0	0.0	0.0	3.8	0.0
AT5G10540.1	0.0	0.0	0.0	0.0	0.0	8.5	0.0	3.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT5G16130.1	0.0	0.0	0.0	0.0	0.0	0.0	4.3	0.0	0.0	0.0	0.0	10.3	0.0	0.0	0.0	0.0
AT5G06600.1	0.0	0.0	0.0	0.0	0.0	0.0	4.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3.2	0.0
AT1G08520.1	0.0	0.0	0.0	11.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3.8	0.0
AT4G25080.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3.1	0.0	0.0	0.0	6.6	0.0	0.0
AT2G47450.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	4.0	3.8	0.0
AT5G17990.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	4.0	3.8	0.0
AT5G49030.1	0.0	0.0	0.0	0.0	0.0	0.0	5.3	0.0	3.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT5G62790.1	0.0	0.0	0.0	0.0	0.0	0.0	3.2	4.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT2G42130.4	0.0	0.0	0.0	14.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.7	0.0	0.0
AT3G18890.1	0.0	0.0	0.0	0.0	0.0	3.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	5.1	0.0
AT4G29410.1	0.0	0.0	0.0	0.0	0.0	0.0	1.1	0.0	0.0	0.0	0.0	20.5	0.0	0.0	0.0	0.0
AT5G27850.1	0.0	0.0	0.0	0.0	0.0	0.0	3.2	0.0	0.0	0.0	12.6	0.0	0.0	0.0	0.0	0.0
AT5G03290.1	0.0	0.0	0.0	0.0	0.0	0.0	2.1	0.0	0.0	0.0	0.0	0.0	0.0	5.3	0.0	0.0

Supplemental Table 2 (continued)

Accession (experiment # in PPDB: 2164 2165 2190 2191 2192 2193 2224 2225 2208 2209 2210 2211 2214 2215 2216 2217)	10000* NadjSPC exp. 2164 P3 control stroma	10000* NadjSPC exp. 2165 P3 S164A stroma	10000* NadjSPC exp. 2190 P3 control stroma	10000* NadjSPC exp. 2191 P3 S164A stroma	10000* NadjSPC exp. 2192 P3 S164A	10000* NadjSPC exp. 2193 P3 control	10000* NadjSPC exp. 2224 P3 control	10000* NadjSPC exp. 2225 P3 S164A	10000* NadjSPC Exp. 2208 P5 control	10000* NadjSPC exp. 2209 P5 S193A	10000* NadjSPC exp. 2210 P5 control	10000* NadjSPC exp. 2211 P5 S193A	10000* NadjSPC exp. 2214 P5 control	10000* NadjSPC exp. 2215 P5 S193A	10000* NadjSPC exp. 2216 P5 control	10000* NadjSPC exp. 2217 P5 S193A
AT1G71040.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	4.0	3.8	0.0
AT1G79920.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.6	9.1
AT3G02520.1	0.0	0.0	0.0	0.0	0.0	0.0	4.9	1.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT2G44530.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.9	6.0	0.0	0.0
AT4G10300.1	0.0	0.0	0.0	0.0	2.4	7.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT2G38140.1	22.8	0.0	20.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT3G49680.1	0.0	0.0	0.0	0.0	0.0	0.0	3.2	2.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT5G10160.1	0.0	0.0	0.0	0.0	2.4	7.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT5G13420.1	0.0	0.0	0.0	0.0	0.0	0.0	3.2	2.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT3G23400.1	0.0	0.0	6.7	14.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT2G09990.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	6.3	0.0	12.3	0.0	0.0	0.0	0.0
AT2G20140.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.0	4.5	0.0
AT4G29040.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.0	4.5	0.0
AT3G22200.2	0.0	0.0	0.0	0.0	7.1	3.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT1G13930.1	0.0	0.0	0.0	0.0	7.1	3.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT1G74920.1	0.0	0.0	0.0	0.0	7.1	3.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT3G15730.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	15.2	0.0	0.0	0.0	0.0	1.3	0.0	0.0
AT5G19820.1	0.0	0.0	0.0	0.0	0.0	0.0	1.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	5.1	0.0
AT5G41670.1	0.0	0.0	0.0	0.0	0.0	0.0	2.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	6.9
AT1G33811.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.7	3.8	0.0
AT2G25280.1	0.0	0.0	0.0	0.0	0.0	0.0	4.3	1.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT2G40290.1	0.0	0.0	0.0	0.0	0.0	0.0	2.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3.8	0.0
AT5G43010.1	0.0	0.0	0.0	0.0	0.0	0.0	3.5	0.0	0.0	0.0	0.0	0.0	2.8	0.0	0.0	0.0
AT5G61170.1	0.0	0.0	0.0	0.0	0.0	0.0	3.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.3	0.0
AT1G16470.1	0.0	0.0	0.0	0.0	0.0	0.0	2.1	3.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT1G79210.1	0.0	0.0	0.0	0.0	0.0	0.0	2.1	3.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT4G25740.1	0.0	0.0	0.0	0.0	0.0	0.0	2.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3.2	0.0
AT5G52650.1	0.0	0.0	0.0	0.0	0.0	0.0	2.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3.2	0.0
AT5G66140.1	0.0	0.0	0.0	0.0	0.0	2.0	0.0	4.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT3G59760.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	12.3	2.2	0.0	0.0	0.0
AT5G04360.1	0.0	0.0	0.0	0.0	0.0	0.0	2.1	0.0	0.0	0.0	0.0	0.0	3.7	0.0	0.0	0.0
AT4G01900.1	0.0	0.0	13.4	7.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT4G39970.1	0.0	0.0	0.0	0.0	4.7	3.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT5G03650.1	0.0	0.0	0.0	0.0	0.0	0.0	2.1	2.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT5G36880.1	0.0	0.0	0.0	7.4	0.0	0.0	2.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT4G25100.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	7.6	6.3	0.0	0.0	0.0	0.0	0.0	0.0
ATCG01110.1	0.0	0.0	0.0	7.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.6	0.0
ATCG00730.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	4.1	0.0	0.0	3.8	0.0
AT5G53490.1	0.0	0.0	0.0	0.0	0.0	5.9	1.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT2G20450.1	0.0	0.0	0.0	0.0	0.0	0.0	3.2	0.0	0.0	0.0	0.0	0.0	1.9	0.0	0.0	0.0
AT2G26060.1	0.0	0.0	0.0	0.0	2.4	0.0	0.0	4.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT4G11820.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	5.6	0.0	1.3	0.0
AT1G08830.1	0.0	0.0	0.0	0.0	0.0	3.9	0.0	2.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT5G37850.1	0.0	0.0	0.0	0.0	0.0	0.0	2.1	2.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT5G54160.1	0.0	0.0	0.0	0.0	0.0	0.0	2.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.6	0.0
AT2G06050.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.8	0.0	0.0	0.0	8.2	0.0	0.0	0.0	0.0
AT5G55480.1	0.0	0.0	0.0	0.0	0.0	5.9	1.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT1G12000.1	0.0	0.0	0.0	0.0	0.0	0.0	1.1	0.0	0.0	0.0	0.0	0.0	0.0	4.0	0.0	0.0
AT1G22840.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	5.6	0.0	0.0	2.3
AT1G79720.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.4	0.0	0.0	0.0	0.0	5.6	0.0	0.0	0.0
AT3G06930.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.9	4.0	0.0	0.0
AT3G13750.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	11.4	3.1	0.0	0.0	0.0	0.0	0.0	0.0
AT3G29090.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.9	4.0	0.0	0.0
AT4G21650.1	0.0	0.0	0.0	0.0	0.0	0.0	2.1	2.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT5G55450.1	0.0	0.0	0.0	0.0	0.0	0.0	2.1	2.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT1G75780.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	5.3	0.0	0.0	0.0	0.0	2.6	0.0
AT1G22530.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	8.4	0.0	2.8	0.0	0.0	0.0
AT4G18360.1	0.0	0.0	0.0	0.0	5.9	2.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT5G65260.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.9	4.6
AT4G14680.1	0.0	0.0	8.7	0.0	0.0	0.0	2.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT1G52360.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.0	2.2	0.0
AT3G15980.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.0	2.2	0.0
AT5G19440.1	0.0	0.0	0.0	0.0	0.0	2.0	0.0	0.0	8.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT1G48420.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	4.1	0.0	0.0	2.6	0.0
AT1G52670.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.4	7.6	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT2G10940.1	0.0	0.0	0.0	0.0	0.0	0.0	1.1	0.0	7.6	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT1G75350.1	11.4	0.0	0.0	0.0	0.0	0.0	2.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT1G24360.1	0.0	0.0	0.0	0.0	0.0	2.0	0.0	2.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT1G63970.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.8	0.0	0.0	0.0	4.1	0.0	0.0	0.0	0.0
AT2G25830.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	7.6	0.0	0.0	0.0	0.0	0.0	0.0	2.3
AT3G14930.1	0.0	0.0	0.0	0.0	0.0	2.0	0.0	2.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT3G48560.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.3	4.6
AT5G13030.1	0.0	0.0	0.0	3.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3.7	0.0	0.0	0.0
AT5G57850.1	0.0	0.0	0.0	0.0	0.0	0.0	1.1	2.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT2G21530.1	0.0	0.0	0.0	0.0	0.0	0.0	2.1	1.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT1G48830.1	0.0	0.0	0.0	0.0	0.0	0.0	2.1	0.0	0.0	0.0	0.0	4.1	0.0	0.0	0.0	0.0
AT2G39460.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	4.7	0.0	6.2	0.0	0.0	0.0	0.0
AT3G55280.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	4.7	0.0	6.2	0.0	0.0	0.0	0.0

Supplemental Table 2 (continued)

Accession (experiment # in PPDB: 2164 2165 2190 2191 2192 2193 2224 2225 2208 2209 2210 2211 2214 2215 2216 2217)	10000* NadjSPC exp. 2164 P3 control stroma	10000* NadjSPC exp. 2165 P3 S164A stroma	10000* NadjSPC exp. 2190 P3 control stroma	10000* NadjSPC exp. 2191 P3 S164A stroma	10000* NadjSPC exp. 2192 P3 S164A	10000* NadjSPC exp. 2193 P3 control	10000* NadjSPC exp. 2224 P3 control	10000* NadjSPC exp. 2225 P3 S164A	10000* NadjSPC Exp. 2208 P5 control	10000* NadjSPC exp. 2209 P5 S193A	10000* NadjSPC exp. 2210 P5 control	10000* NadjSPC exp. 2211 P5 S193A	10000* NadjSPC exp. 2214 P5 control	10000* NadjSPC exp. 2215 P5 S193A	10000* NadjSPC exp. 2216 P5 control	10000* NadjSPC exp. 2217 P5 S193A
AT5G48230.1	0.0	0.0	0.0	0.0	0.0	0.0	2.1	1.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT1G59900.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.7	1.3	0.0
AT1G78850.1	0.0	0.0	0.0	0.0	4.7	2.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT2G44160.1	0.0	0.0	0.0	0.0	0.0	0.0	1.1	0.0	0.0	0.0	0.0	0.0	0.0	2.7	0.0	0.0
AT4G16660.1	0.0	0.0	0.0	0.0	0.0	2.0	2.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT5G07350.1	0.0	0.0	0.0	0.0	0.0	3.9	1.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT1G06000.1	0.0	0.0	0.0	0.0	0.0	2.0	0.0	2.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT1G76680.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.8	0.0	3.1	0.0	0.0	0.0	0.0	0.0	0.0
AT1G79690.1	0.0	0.0	0.0	0.0	0.0	3.9	0.0	0.0	3.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT2G44210.1	0.0	0.0	0.0	0.0	0.0	0.0	2.1	0.0	0.0	0.0	0.0	0.0	1.9	0.0	0.0	0.0
AT3G07660.1	0.0	0.0	0.0	0.0	0.0	2.0	2.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT3G08900.1	0.0	0.0	0.0	0.0	4.0	2.6	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT3G16410.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.4	0.0	0.0	0.0	0.0	3.7	0.0	0.0	0.0
AT3G24170.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.4	0.0	0.0	0.0	8.2	0.0	0.0	0.0	0.0
AT4G27320.1	0.0	0.0	0.0	0.0	0.0	0.0	2.1	1.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT5G23860.1	0.0	0.0	0.0	0.0	0.0	0.0	1.7	0.0	0.0	3.8	0.0	0.0	0.0	0.0	0.0	0.0
AT1G77510.1	0.0	0.0	0.0	0.0	3.5	2.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT5G26780.1	0.0	0.0	0.0	0.0	3.5	2.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT3G46040.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.8	0.0	0.0	0.0	0.0	0.0	0.0	1.3	0.0
AT4G22930.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.4	3.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT3G54210.1	0.0	0.0	0.0	0.0	0.0	0.0	1.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.3	0.0
ATCG00650.1	0.0	0.0	0.0	0.0	0.0	0.0	1.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.3	0.0
AT1G18500.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.3
AT1G70820.1	0.0	0.0	6.7	0.0	0.0	2.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT4G25050.1	0.0	0.0	0.0	0.0	2.4	2.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT5G22800.1	0.0	0.0	0.0	3.7	0.0	0.0	0.0	0.0	0.0	3.1	0.0	0.0	0.0	0.0	0.0	0.0
AT2G26540.1	0.0	0.0	6.7	3.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT2G27020.1	0.0	0.0	0.0	0.0	2.4	0.0	0.0	1.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT5G23250.1	0.0	0.0	0.0	0.0	0.0	1.0	0.0	2.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT1G75280.1	0.0	0.0	0.0	0.0	0.0	2.0	0.0	0.0	3.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT5G53480.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	4.2	0.0	0.0	0.0	1.3	0.0
AT3G01910.1	0.0	0.0	0.0	0.0	0.0	0.0	1.1	0.0	0.0	0.0	0.0	4.1	0.0	0.0	0.0	0.0
AT3G42050.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.3	1.3	0.0
AT1G11580.1	0.0	0.0	0.0	0.0	0.0	0.0	1.1	1.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT1G53580.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.4	3.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT1G73620.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.4	0.0	0.0	0.0	0.0	1.9	0.0	0.0	0.0
AT1G75040.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.4	0.0	0.0	0.0	4.1	0.0	0.0	0.0	0.0
AT2G42500.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.9	0.0	1.3	0.0
AT3G02760.1	0.0	0.0	0.0	0.0	0.0	2.0	1.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT4G09580.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3.8	0.0	0.0	0.0	0.0	1.3	0.0	0.0
AT4G34980.1	0.0	0.0	0.0	0.0	0.0	0.0	1.1	1.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT4G35800.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.3	1.3	0.0
AT5G34850.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.4	0.0	0.0	0.0	4.1	0.0	0.0	0.0	0.0
AT5G49810.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.3	2.3
AT5G55810.1	0.0	0.0	0.0	0.0	0.0	0.0	1.1	1.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
AT5G58190.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.3	1.3	0.0
AT5G64080.1	0.0	0.0	0.0	0.0	0.0	2.0	0.0	1.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

cloned into pCR™8/GW/TOPO®vector (Invitrogen) and binary vector pMDC123 using Gateway® LR Clonase® II enzyme mix (Invitrogen) and then introduced into the *Agrobacterium tumefaciens* GV3101 by electroporation. Agrobacterium-mediated transformation by flower dip methods was according to (*47*). Crude plant genomic DNA was extracted from grinding frozen leaf tissues in the gDNA extraction buffer (0.2 M Tris-HCl, pH 7.5, 0.25 M NaCl, 0.025 M EDTA, pH 8.0, 0.5% SDS) and precipitated by isopropanol. Genotyping was performed using *CLPP3/5*-specific (primer set 4), transgene-specific primers (primer set 5), and T-DNA specific primers (primer set 6) Total RNA was extracted from grinding frozen leaf tissues using RNeasy Plant Mini Kit (Qiagen). The first-strand of cDNA was synthesized using SuperScript® III First-Strand Synthesis System (Invitrogen) and then used for the synthesis of the second-strand cDNA. Primers are listed in Supplemental Table 1.

3.4.2 Plant growth and phenotypic analysis

WT, *clpp3-1* null, and complemented CLPP3-STREPII and CLPP3S164A-STREPII T2 seeds were surface disinfected with 70% ethanol for 15 min and then rinsed with 95% ethanol for three times before sowing on the half Murashige and Skoog (1/2 MS) plates. These plates were put in a cold room (4°C, dark) for 3 days stratification and then transferred into the growth chamber with 16/8 light/dark at 70 µmol photons/m^2s for 12 days. Plates were relocated and observed every day. For soil-grown phenotypic analysis, wt, complemented *CLPP3-STREPII* and *CLPP3S164A-STREPII* T3 seeds were sown on soil, put in a cold room for 3 days stratification, and then transferred in a growth chamber (16/8 light /dark at 120-150 µmol photons/m^2s). These soil-growing plants were observed every two days. The plant height and the

number of the rosettes leaves were recorded for the growth and development rate. Complemented CLPP5-STREPII and transgenic CLPP5S193A-STREPII were grown side by side on soil after 5 days stratification and then grown under 10/14 light/dark at 120-150 µmol photons/m^2s. BASTA-spraying was performed every 2-3 days for the first 14 days. Survival CLPP5-STREPII and CLPP5S193A-STREPII were genotyped and observed every week. At least ten developing siliques with similar size (length 1.8-2.0 cm) from heterozygous *clpp5-1* and heterozygous *clpp5-1* with *CLPP5S193A-STREPII* T2 plants were exanimated under the dissecting microscope. Seeds were counted according to its colors.

3.4.3 SDS-PAGE, BN-PAGE, and immunoblot

Total leaf protein extraction under denaturing condition and non-denaturing condition was performed according to (*48*) and (*15*), respectively. Isolation of stromal proteins was according to (*17*). For immunoblot analysis of transgene expression and the assembly state of CLPPRT complexes, 20 µg of total leaf soluble proteins or 10-50 µg of stromal proteins were separated on SDS-PAGE or Novex pre-cast Bis-Tris 4-16% gel (Invitrogen), transferred onto nitrocellulose membranes or PVDF membranes, stained in Ponceau S solution (0.3% Ponceau in 3% TCA). The immunoblot blot using anti-STREPII, anti-CLPR2 and, anti-CLPP3 antiserum was according to (*17*).

3.4.4 Affinity purification of STREPII-tagged CLP complexes and MS/MS analysis

At least 2 mg stromal proteins or 24 mg total leaf proteins extracted under the non-denaturing conditions was loaded on a self-packed StrepTactin column using the superflow high capacity resin (IBA) according to (*15*), except that 5 mM biotin instead

of 2.5 mM desthiobiotin was included in the elution step. To remove biotin-conjugated proteins founded in the affinity eluates, avidin (0.2 mg per g leaf tissue) was later included in the extraction buffer during total soluble protein extraction from CLPP5-STREPII and CLPP5S193A-STRERPII lines. Amicon and Microcon spin concentrators (3 kDa) were used for concentration of the eluates before further analysis.

Concentrated eluates were separated on 10.5-14% precast gels (Biorad), followed by the MS-compatible silver stain. Comparative proteome analysis using an LTQ-Orbitrap mass spectrometer, data processing, database searches, quantification of the relative protein abundance, as well as the selection of the best gene models followed the procedures described in Friso (2011). We evaluated the samples for potential enrichment based on matched MS/MS adjusted spectra (adjSPC) normalized to the total number of adjSPC in each sample, resulting in NadjSPC. Annotations are from the Plant Proteome Data Base (http://ppdb.tc.cornell.edu/).

REFERENCES

1. R. T. Sauer, T. A. Baker, AAA+ proteases: ATP-fueled machines of protein destruction. *Annual Review of Biochemistry* **80**, 587-612 (2011).
2. J. A. Alexopoulos, A. Guarne, J. Ortega, ClpP: a structurally dynamic protease regulated by AAA+ proteins. *Journal of Structural Biology* **179**, 202-210 (2012).
3. K. Liu, A. Ologbenla, W. A. Houry, Dynamics of the ClpP serine protease: a model for self-compartmentalized proteases. *Critical Reviews in Biochemistry and Molecular Biology* **49**, 400-412 (2014).
4. K. Nishimura, K. J. van Wijk, Organization, function and substrates of the essential Clp protease system in plastids. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **1847**, 915-930 (2015).
5. K. Nishimura, Y. Kato, W. Sakamoto, Essentials of proteolytic machineries in chloroplasts. *Mol Plant* **10**, 4-19 (2017).
6. S. A. Mahmoud, P. Chien, Regulated proteolysis in bacteria. *Annual Review of Biochemistry* **87**, 677-696 (2018).
7. N. J. Kuhlmann, P. Chien, Selective adaptor dependent protein degradation in bacteria. *Current Opinion in Microbiology* **36**, 118-127 (2017).
8. A. O. Olivares, T. A. Baker, R. T. Sauer, Mechanical protein unfolding and degradation. *Annual Review of Physiology* **80**, 413-429 (2018).
9. J. B. Peltier *et al.*, Clp protease complexes from photosynthetic and non-photosynthetic plastids and mitochondria of plants, their predicted three-dimensional structures, and functional implications. *Journal of Biological Chemistry* **279**, 4768-4781 (2004).
10. J. Wang, J. A. Hartling, J. M. Flanagan, The structure of ClpP at 2.3 A resolution suggests a model for ATP-dependent proteolysis. *Cell* **91**, 447-456 (1997).

11. J. Schelin, F. Lindmark, A. K. Clarke, The clpP multigene family for the ATP-dependent Clp protease in the cyanobacterium Synechococcus. *Microbiology* **148**, 2255-2265. (2002).
12. T. M. Stanne, E. Pojidaeva, F. I. Andersson, A. K. Clarke, Distinctive types of ATP-dependent Clp proteases in cyanobacteria. *Journal of Biological Chemistry* **282**, 14394-14402 (2007).
13. F. I. Andersson *et al.*, Structure and function of a novel type of ATP-dependent Clp protease. *Journal of Biological Chemistry* **284**, 13519-13532 (2009).
14. M. El Bakkouri *et al.*, Structural insights into the inactive subunit of the apicoplast-localized caseinolytic protease complex of Plasmodium falciparum. *Journal of Biological Chemistry* **288**, 1022-1031 (2013).
15. P. D. Olinares, J. Kim, J. I. Davis, K. J. van Wijk, Subunit stoichiometry, evolution, and functional implications of an asymmetric plant plastid ClpP/R protease complex in Arabidopsis. *Plant Cell* **23**, 2348-2361 (2011).
16. A. Y. Yu, W. A. Houry, ClpP: a distinctive family of cylindrical energy-dependent serine proteases. *FEBS Letters* **581**, 3749-3757 (2007).
17. J. Kim *et al.*, Structures, functions, and interactions of ClpT1 and ClpT2 in the Clp protease system of Arabidopsis chloroplasts. *Plant Cell* **27**, 1477-1496 (2015).
18. L. L. Sjogren, A. K. Clarke, Assembly of the chloroplast ATP-dependent Clp protease in Arabidopsis is regulated by the ClpT accessory proteins. *Plant Cell* **23**, 322-332 (2011).
19. J. C. Moreno *et al.*, Generation and characterization of a collection of knock-down lines for the chloroplast Clp protease complex in tobacco. *Journal of Experimental Botany* **68**, 2199-2218 (2017).
20. J. Kim *et al.*, Modified Clp protease complex in the ClpP3 null mutant and consequences for chloroplast development and function in Arabidopsis. *Plant Physiology* **162**, 157-179 (2013).

21. J. Kim *et al.*, Subunits of the plastid ClpPR protease complex have differential contributions to embryogenesis, plastid biogenesis, and plant development in Arabidopsis. *Plant Cell* **21**, 1669-1692 (2009).
22. A. Rudella, G. Friso, J. M. Alonso, J. R. Ecker, K. J. van Wijk, Downregulation of ClpR2 leads to reduced accumulation of the ClpPRS protease complex and defects in chloroplast biogenesis in Arabidopsis. *Plant Cell* **18**, 1704-1721 (2006).
23. S. Koussevitzky *et al.*, An Arabidopsis thaliana virescent mutant reveals a role for ClpR1 in plastid development. *Plant Molecular Biology* **63**, 85-96 (2007).
24. B. Zheng, T. M. MacDonald, S. Sutinen, V. Hurry, A. K. Clarke, A nuclear-encoded ClpP subunit of the chloroplast ATP-dependent Clp protease is essential for early development in Arabidopsis thaliana. *Planta* **224**, 1103-1115 (2006).
25. K. Nishimura *et al.*, ClpS1 is a conserved substrate selector for the chloroplast Clp protease system in Arabidopsis. *Plant Cell* **25**, 2276-2301 (2013).
26. J. Apitz *et al.*, Posttranslational control of ALA synthesis includes GluTR degradation by Clp protease and stabilization by GluTR-binding protein. *Plant Physiology* **170**, 2040-2051 (2016).
27. W. Tapken, J. Kim, K. Nishimura, K. J. van Wijk, M. Pilon, The Clp protease system is required for copper ion-dependent turnover of the PAA2/HMA8 copper transporter in chloroplasts. *New Phytologist* **205**, 511-517 (2015).
28. R. Welsch *et al.*, Clp protease and OR directly control the proteostasis of phytoene synthase, the crucial enzyme for carotenoid biosynthesis in Arabidopsis. *Mol Plant* **11**, 149-162 (2018).
29. P. Pulido *et al.*, Specific Hsp100 chaperones determine the fate of the first enzyme of the plastidial isoprenoid pathway for either refolding or degradation by the stromal Clp protease in Arabidopsis. *PLoS Genetics* **12**, e1005824 (2016).

30. K. Nishimura, Y. Kato, W. Sakamoto, Chloroplast proteases: updates on proteolysis within and across suborganellar compartments. *Plant Physiology* **171**, 2280-2293 (2016).
31. Y. Wu *et al.*, Quantitative proteomic analysis of two different rice varieties reveals that drought tolerance is correlated with reduced abundance of photosynthetic machinery and increased abundance of ClpD1 protease. *Journal of Proteomics* **143**, 73-82 (2016).
32. J. C. Moreno *et al.*, Temporal proteomics of inducible RNAi lines of Clp protease subunits identifies putative protease substrates. *Plant Physiology* **176**, 1485-1508 (2018).
33. J. M. Flynn, S. B. Neher, Y. I. Kim, R. T. Sauer, T. A. Baker, Proteomic discovery of cellular substrates of the ClpXP protease reveals five classes of ClpX-recognition signals. *Molecular Cell* **11**, 671-683. (2003).
34. J. Feng *et al.*, Trapping and proteomic identification of cellular Substrates of the ClpP protease in Staphylococcus aureus. *Journal of Proteome Research* **12**, 547-558 (2013).
35. D. B. Trentini *et al.*, Arginine phosphorylation marks proteins for degradation by a Clp protease. *Nature* **539**, 48-+ (2016).
36. F. Fischer, J. D. Langer, H. D. Osiewacz, Identification of potential mitochondrial CLPXP protease interactors and substrates suggests its central role in energy metabolism. *Scientific Reports* **5**, 18375 (2015).
37. D. Y. Kim, K. K. Kim, The structural basis for the activation and peptide recognition of bacterial ClpP. *Journal of Molecular Biology* **379**, 760-771 (2008).
38. J. M. Flynn, I. Levchenko, R. T. Sauer, T. A. Baker, Modulating substrate choice: the SspB adaptor delivers a regulator of the extracytoplasmic-stress response to the AAA+ protease ClpXP for degradation. *Genes & Development* **18**, 2292-2301 (2004).

39. R. M. Raju *et al.*, Mycobacterium tuberculosis ClpP1 and ClpP2 function together in protein degradation and are required for viability in vitro and during infection. *PLoS Pathogens* **8**, e1002511 (2012).

40. B. J. Nikolau, J. B. Ohlrogge, E. S. Wurtele, Plant biotin-containing carboxylases. *Archives of Biochemistry and Biophysics* **414**, 211-222 (2003).

41. K. Nishimura *et al.*, Discovery of a unique Clp component, ClpF, in chloroplasts: a proposed binary ClpF-ClpS1 adaptor complex functions in substrate recognition and delivery. *Plant Cell* **27**, 2677-2691 (2015).

42. J. Arends *et al.*, An integrated proteomic approach uncovers novel substrates and functions of the Lon protease in Escherichia coli. *Proteomics* **18**, (2018).

43. J. Arends, N. Thomanek, K. Kuhlmann, K. Marcus, F. Narberhaus, In vivo trapping of FtsH substrates by label-free quantitative proteomics. *Proteomics* **16**, 3161-3172 (2016).

44. A. K. Livingston, J. A. Cruz, K. Kohzuma, A. Dhingra, D. M. Kramer, An Arabidopsis mutant with high cyclic electron flow around photosystem I (hcef) involving the NADPH dehydrogenase complex. *Plant Cell* **22**, 221-233 (2010).

45. J. C. Moreno Beltran *et al.*, Temporal proteomics of inducible RNAi lines of Clp protease subunits identifies putative protease substrates. *Plant Physiology*, (2017).

46. J. W. Graham, M. G. Lei, C. Y. Lee, Trapping and identification of cellular substrates of the Staphylococcus aureus ClpC chaperone. *Journal of Bacteriology* **195**, 4506-4516 (2013).

47. S. J. Clough, A. F. Bent, Floral dip: a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. *Plant Journal* **16**, 735-743. (1998).

48. G. Friso, P. D. B. Olinares, K. J. van Wijk, *Chloroplast Research in Arabidopsis,* R. P. Jarvis, Ed. (Humana Press, New York, 2011), vol. 775, pp. 265-282.

CHAPTER FOUR

INVESTIGATION OF THE SPATIAL PROXIMITY OF THE CLP SUBUNITS WITHIN THE CHLOROPLAST CLP MACHINERY[1]

4.1 ABSTRACT

Chloroplast CLP proteases have a far greater complexity than other known CLP protease in mitochondria or prokaryotes. It is composed of 10 different subunits (CLPP1, P3-6, R1-4), organized as two heptameric P and R rings, and two accessory proteins CLPT1-2 specific to higher plants. Although the stoichiometry of the CLPPR core is solved, the functions of the unique C-terminal extension of plant CLPPR subunits and the structure of chloroplast CLPPRT are unknown. The CLP protease complex must interact with the CLPC1,2 and CLPD AAA+ chaperones (likely organized as hexameric rings) for substrate selection, unfolding and delivery into the protease core for degradation. Yet, little is known about this interaction between CLP protease and chaperones, nor is it known if the hexameric rings are homo- or heterooligomers. To better understand these structure-function questions, we applied a mass spectrometry (MS) cleavable DSSO crosslinker to the affinity-enriched CLP complexes and investigated the spatial proximity of the CLPP/R/T/C subunits through MS analysis; this approach is termed XL-MS. We obtained XL-MS results from multiple independent experiments using either CLPR4, CLPP3, CLPP5 or CLPC1-TRAP as bait. This chapter summarizes these preliminary results, as well as concepts,

challenges and follow-up analysis, including application of the proximity data for structural modeling and co-evolution analysis.

---

[1]This is an ongoing project with Dr. Clinton Yu and Professor Lan Huang at the Department of Physiology & Biophysics, University of California, Irvine, CA. Jui-Yun prepared chloroplast CLP complexes and performed peptide digestion with the aid from Dr. Giulia Friso. Dr. Clinton Yu conducted the MSn and data analysis.

4.2 INTRODUCTION

Arabidopsis CLP protease is a ~350 kDa protein complex, composed by two heptameric rings, the P-ring and R ring (*1, 2*). Compared to known CLP homologs in mitochondria and non-plant species, chloroplast CLP complexes have several unique features. First, two accessory proteins CLPT1/T2 are associated with CLP complexes (*3, 4*). These two accessory proteins are unique to the plants and are suggested to stabilize the CLPPR complexes through their MYFF motif interacting with a hydrophobic region of the CLPPR complex (*4*); however, the docking site of CLPT1/T2 is unknown. Second, CLPP/R subunits have an additional C-terminal extension as compared to bacterial homologs. These C-terminal extensions are only present in high plants and their function is unclear. Third, chloroplast CLP complexes have a hetero-oligomeric composition with a unique stoichiometry P1: R1: R2: R3: R4: P3: P4: P5: P6= 3: 1: 1: 1: 1: 1: 2: 3: 1 (*2*). These CLPP/R subunits show no or very little (ClpR1 and ClpR3) redundancy to each other and have only 20-40% identity between them (*5*). It is mystery why the chloroplast CLP evolved such

complexity and it is not clear if this represents an adaptation to the plastid/chloroplast proteome. The spatial arrangement of these CLPPR subunits is unknown, and may hold the key to understanding this increased complexity and interaction with CLPT1/2, CLPC1/2, and CLPD chaperones.

Since protein function and structure are closely related, investigation of the protein-protein interactions (PPI) within the complexes not only provides assembly and structural information, but also helps to elucidate possible regulatory mechanisms for protein functions. Current techniques in structural biology include X-ray crystallography, NMR, and cryo-EM. However, technical barriers (difficult for crystallization) and limitations (suitable for small proteins or low resolution) make application of these techniques to study protein complexes in details challenging and difficult (*6*-*8*). For example, preparation of the pure and sufficient amounts of crystallized protein complexes is the bottleneck for X-ray crystallography. Recombinant proteins are commonly used for sample preparation. However, this strategy requires several trials to find the best condition to maximize the protein yield. The chloroplast CLP complex has a complicated composition that makes reconstitution from recombinant proteins challenging and it is not clear if this would yield complexes that represent the native state. NMR is limited to proteins with a small size (< 50kDa), which does not work for CLPPRT complexes. Cryo-EM in increasingly successful at obtained high resolution structures and the chloroplast CLP system would be a very suitable target, assuming we can obtain a sufficiently enriched solution of native complexes.

Recently, a XL-MS approach (MS analysis of crosslinked peptides) has become an emerging tool to study the potential PPI within a complex through the residue-residue proximity analysis restrained by the length of crosslinkers (*9, 10*). This approach is complementary to the techniques mentioned above and can capture critical information about inter- and intra-protein contact sites.

So far, little structural information is available for chloroplast CLPPRT complexes. Here we conducted XL-MS analysis of the chloroplast CLP complexes using a symmetric MS-cleavable crosslinker DSSO (disuccinimidyl sulfoxide) and a highly sensitive MS instrument (Figure 1). DSSO is lysine-reactive crosslinker with 10-12 angstrom armlength. Due to its symmetric MS-cleavable nature, three types of the links (inter-subunit, intra-subunit, and dead end) can be distinguished at MS2 through their unique fragment pattern in the mass spectrometer. This facilitates the interpretation of MS data of protein-protein proximity (inter-subunit links) and provides structural details for protein complexes (*11*). Our goal of the XL-MS analysis is to prepare and collect the residue-residue proximity data for mapping the 3D structure and understand the regulation mechanisms of the chloroplast CLP complex.

4. 3 MATERIALS AND METHODS

4.3.1 Affinity Purification of CLPRT cores using total leaf tissues

To enrich CLP complexes for *in vitro* crosslinking, different strep-tagged transgenic lines CLPP5-strepII and cCLPR4-strepII were used to enrich CLP complexes P-ring, R-ring, and CLPPR cores, using affinity protocols as described in (*2, 4*). Since CLPPRTs copurified with strep-tagged CLPC1E374AE718A, this

CLPC1E374AE718A line was also used to purify CLP complexes, with the aim to capture the transient interaction between CLPC1 and CLPPRT *in vivo*.

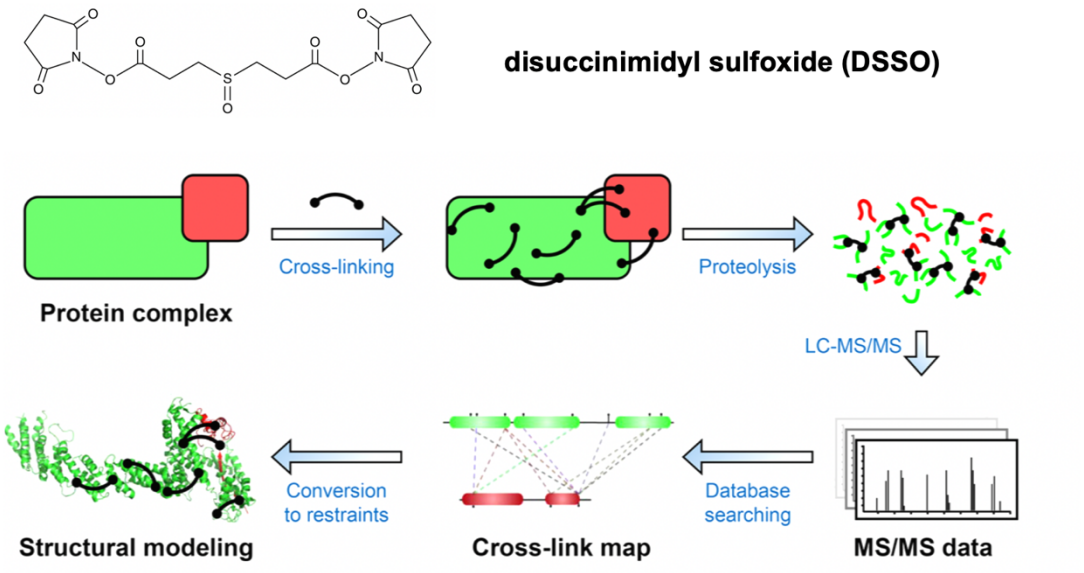


Figure 1. The workflow of XL-MS.
Adapted with permission from Yu and Huang et al (*12*) ©2018, American Chemical Society.

For Clp purification, rosette leaves were ground into fine powder in liquid nitrogen and resuspend with extraction buffer (with Pefabloc SC and avidin). After three time mixing by vortex, the leaf extract was filtered with 3-4 layers of miracloth, followed by ultracentrifuge at 100k g at least 1 hr at 4°C. The supernatant (in some cases after concentration using 50 kDa Amicon concentrators) was loaded on 1 ml strep columns (IBA, StrepTactinXT). The affinity purification followed the

manufactural protocol except that additional 5 times wash steps were included. Eluate fractions were then concentrated using 50 kDa Amicon concentrator before use.

4.3.2 DSSO crosslinking of affinity purified CLP complexes

Various DSSO concentration (0, 0.25, 0.75, 1.0 mM) were tested for crosslinking of the affinity elute products to determine the optimal concentration of DSSO crosslinking at given protein concentrations. The most optimal concentration was used in subsequent larger scale cross-linking experiments.

4.3.3 Sample preparation and MSn analysis

Affinity-enriched CLP complexes were concentrated using Microcon-30 (30 kDa cutoff). Microcons were washed with 300 µl of milliQ water by spinning at 14k g for 5 min before use. The protein sample was added to the Microcon and span at 14k g for 5 min or more for concentration. The concentrated sample was washed with 300 µl of 20 mM Hepes by spinning at 14k g for 25 min. For sample denaturation, 450 µl of 8M urea dissolved in 25 mM ammonium bicarbonate was added. Sample reduction was performed by adding 2 µl of 100 mM DTT dissolved in 25 mM ammonium bicarbonate and incubating at RT for 30 min. For sample alkylation, 5 µl of 100 mM iodoacetamide (prepare freshly) was added. After 30 min RT incubation for denaturation and alkylation steps, the sample was collected by spinning at 14k g for 5 min and washed twice with 300 µl Hepes by spinning at 14k g for 25 min. For additional denaturation, 20 µl of 8M urea in 25 mM ammonium bicarbonate was added. Adjust the pH of the sample solution by adding additional 85 µl of 25 mM ammonium bicarbonate (the end urea concentration is 1.5 M) before trypsin digestion. Trypsin digestion was performed using trypsin solution (protease: protein mass ratio=

1: 20) and incubated at 37°C for 14 hr. The digested sample was collected by spinning at 14k g for 25 min. The peptide sample collected in the flow-through was desalted using Pierce C18 column according to the manufacture manual. Desalted peptides were dried and resuspended in 10 µl 2% acetone and 1% formaldehyde. MS analysis was performed using the Thermo Scientific Orbitrap Fusion Lumos Tribrid Mass Spectrometer at the proteomic facility, University of California Irvine according to (*11*).

4.4 RESULT AND DISCUSSION

4.4.1 In vitro crosslinking of CLPPRT complexes

To investigate the proximity of the CLP subunits within the tetradecamer, we first tested *in vitro* crosslinking of the affinity eluates pulled down from strep-tagged CLP lines. We decided to carry out the crosslinking after affinity-enrichment of the Clp complexes, because previous experiments showed that the purified Clp complex was relative stable (*2*). The alternative approach is to cross-link soluble leaf protein extracts or chloroplast protein extracts but that would require more DSSO crosslinker and could result in protein degradation during the incubation at room temperature with DSSO. Such potential and likely uncontrolled degradation is less of a problem when Clp complexes are first purified.

4.4.2 Titration of DSSO concentration for crosslinking

To optimize the crosslinking reaction, we first performed a titration test to determine the optimal concentration for DSSO crosslinking using affinity-enriched CLP product extracted from the strep-tagged CLP lines. Since the yield of CLP in the

affinity eluates from 10 g tissues was too limited for quantification by protein concentration essays (e.g. Bradford), we estimated the yield through western blot signals with several assumptions based on previous proteome analysis of the transgenic CLP lines. These assumptions are that i) stromal proteins account for about half of the total soluble leaf proteins, ii) CLP complexes are about 0.01% of stroma proteins (*13*), and iii) the affinity purification yield of CLP is 5% (based on prior experience). After applying these factors, we could calculate the approximate amount of the CLPPRT complexes in the affinity eluates for subsequent crosslinking. For example, from 20 g leaf tissue, we could extract 160 mg of total soluble leaf proteins (quantified using BCA assay), resulting in an estimated 0.008 mg Clp core protein. Considering the 5% purification efficiency, the estimate yield of CLP in the affinity eluates is ~0.4 μg of CLPPRT (350 kDa) in the affinity eluate (~1 pmol of complex).

To find the optimal DSSO concentration, we applied different concentrations of DSSO (0, 0.5, 0.75, or 1.0 mM) to the affinity eluates from the strep-tagged CLP lines and incubated them at room temperature for 30 min. In the previously published protocols for other non-plant complexes (*11, 14, 15*), incubations were done at 37°C but we wanted to avoid such non-physiological temperatures. Affinity eluates from before and after crosslinking were loaded on the same gel for western blot analysis against strepII antiserum. The optimal concentration was determined based on higher molecular mass products and loss of monomeric signals (Figure 2A-B).

4.4.3 DSSO crosslinking and MS analysis

To study the protein-protein interactions within the CLPPRT complexes, several different strep-tagged lines were used for CLP purification and DSSO

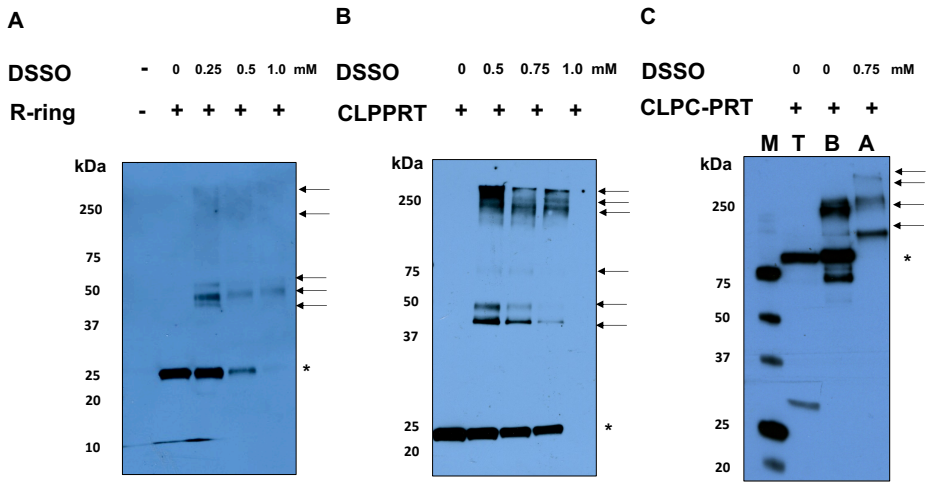


Figure 2. Western blot of the affinity eluates before and after DSSO crosslinking.

A, Titration of DSSO using the affinity eluates from the strep-tagged CLPR4 lines. The protein input for DSSO titration is the CLP R ring. Anti-strepII antiserum was used for western blot. *: strep-tagged CLPR4. Arrows: higher molecular mass product. B, Titration of DSSO using the affinity eluates from the strep-tagged CLPP5 lines. The protein input for DSSO titration is the CLP complexes purified from the strep-tagged CLPP5 lines. Anti-strepII antiserum was used for western blot. *: strep-tagged CLPP5. Arrows: higher molecular mass product. C, DSSO crosslinking of the affinity eluates from the strep-tagged CLPC1E374AE718A lines. The protein input for DSSO crosslinking is the CLP complexes purified from the strep-tagged CLPC1E374AE718A lines. M: marker; B: before XL; A: after XL. Anti-strepII antiserum was used for western blot. *: strep-tagged CLPCE374AE718A. Arrows: higher molecular mass product.

crosslinking. These lines include the complemented strep lines CLPR4 (*clpr4-1*/CLPR4-STREPII), CLPP3 (clpp3-1/CLPP5-STREPII) and CLPP5 (*clpp5-1*/CLPP5-STREPII). These lines were chosen because of the location of the strep-tagged subunit (in P or R ring) and their (nearly) fully complemented phenotypes. Previously, we could still detect the individual P/R-ring in addition to the whole CLPPR cores in the eluates by BN-PAGE and mass spectrometry (chapter three). We therefore used these lines as different baits to enrich the CLPPR cores and P/R-rings.

We first tested DSSO crosslinking using the affinity eluates extracted from 20 g strep-tagged CLPP3 or CLPP5 plants. XL-MS of these two samples gave several inter-subunit links, including CLPR4-CLPP4 (between R-ring and P-ring), CLPR3-CLPP3 (between R-ring and P-ring), CLPR4-CLPR2 (within R ring). No inter-subunit links were found within the P ring, and this is likely because of the relative low number of lysine residues (where the DSSO reacts with) in CLPP subunits as compared to CLPR subunits (see (*16*)). Known endogenous contaminants (streptavidin column interactors such as biotin carboxylase) and a few non-plastid proteins were also observed in samples, but they did not link to CLP subunits.

To increase the coverage and reduce the cellular contaminants for streptavidin resin, we doubled the amount of starting leaf material 40 g (~0.8 μg Clp core) and treated the total soluble proteins with avidin (0.2 mg per g tissues) prior to affinity purification to reduce biotin-conjugated protein contaminants. Since CLPP3 and CLPP5 are both located in the CLPP ring and CLPP5 has more copy number (3x) in the tetradecamer than CLPP3 (1x), we used the strep-tagged CLPP5 rather than CLPP3, in addition to strep-tagged CLPR4 for the second XL-MS analysis. A total 32

different inter-subunit links were identified. In the CLPR4 sample, 22 unique inter-subunit links were identified. In the CLPP5 sample, 20 unique inter-subunit links were identified (Table 1). Only 10 links were identified in both CLPP5 and CLPR4 samples. In addition, the inter-subunit links within the P ring were only found in CLPP5 samples but not in CLPR4 samples, suggesting that using the affinity product extracted from the different strep-tagged CLP lines can increase the coverage of the crosslinks. The inter-subunit links were summarized in the matrix (Figure 3). Several other chloroplast proteins, not considered part of the CLP system, were also identified: CPN60 was identified in CLPR4 sample, while Hsp70 were found in both CLPP5 and CLPR4 samples, but there is no indication that they directly interacted with the CLP complex. These chaperones are abundant *in vivo* (*17*).

Our initial goal for XL-MS analysis is to directly map the spatial arrangement of the CLPPRT complexes through the subunit proximity within the complexes. However, the residue-residue proximity data collected so far is not sufficient to reject any potential spatial arrangement model. It is because most of the inter-subunit links have one end residue in the N/C-terminal or handle regions that are either not present or flexible (conformation change), predicted according to the bacterial CLP homologs (Figure 4 and Table 2). It is difficult to conclude their possible spatial arrangement without further structural information and computational modeling. Additional structural modeling with various types of the spatial arrangement restrained by the residue-residue proximity data from the XL-MS may help to solve the spatial arrangement of CLPPRT complexes. Recently, a new computational approach that

predicts the protein 3D structure by the sequence variation during evolution has been developed (*19, 20*). The idea for this approach is that the variation of the residues

Table 1. Summary of the interlinks among CLP subunits identified through the XL-MS of the eluates from strep-tagged CLPP5/R4 lines

Inter-subunit links	CLPR4	CLPP5
ClpP1 (R-ring):K164-ClpR2 (R-ring):K266	x	
ClpP1 (R-ring):K164-ClpR3 (R-ring):K254	x	x
ClpP1 (R-ring):K7-ClpR1 (R-ring):K114	x	x
ClpP1 (R-ring):K7-ClpR4 (R-ring):K282	x	
ClpP3 (P-ring):K198-ClpP4 (P-ring):K276		x
ClpP3 (P-ring):K198-ClpP4 (P-ring):K276|K280		x
ClpP3 (P-ring):K198-ClpP4 (P-ring):K280		x
ClpP3 (P-ring):K215;K218-ClpP5 (P-ring):K273		x
ClpP3 (P-ring):K277-ClpP6 (P-ring):K241		x
ClpP3 (P-ring):K283;K288;K289-ClpR3 (R-ring):K294	x	
ClpP3 (P-ring):K288;K289-ClpR3 (R-ring):K294	x	
ClpP3 (P-ring):K288|K289-ClpR3 (R-ring):K292	x	
ClpP3 (P-ring):K288|K289-ClpR3 (R-ring):K294		x
ClpP3 (P-ring):K289-ClpP6 (P-ring):K241		x
ClpP3 (P-ring):K289-ClpR3 (R-ring):K292	x	x
ClpP3 (P-ring):K289-ClpR3 (R-ring):K294	x	x
ClpP4 (P-ring):K201-ClpR2 (R-ring):K211	x	x
ClpP4 (P-ring):K227-ClpR2 (R-ring):K211	x	
ClpP4 (P-ring):K227-ClpR3 (R-ring):K254	x	x
ClpP4 (P-ring):K227-ClpR4 (R-ring):K264	x	x
ClpP4 (P-ring):K227-ClpT2 (wP-ring):K158	x	x
ClpP4 (P-ring):K276-ClpP5 (P-ring):K276		x
ClpP6 (P-ring):K241-ClpT1 (wP-ring):K139		x
ClpP6 (P-ring):K241-ClpT1 (wP-ring):K150	x	
ClpR1 (R-ring):K114-ClpR2 (R-ring):K211	x	
ClpR2 (R-ring):K211-ClpR3 (R-ring):K292		x
ClpR2 (R-ring):K211-ClpR4 (R-ring):K217	x	x
ClpR2 (R-ring):K211-ClpR4 (R-ring):K264	x	x
ClpR2 (R-ring):K211-ClpR4 (R-ring):K267	x	
ClpR2 (R-ring):K211-ClpR4 (R-ring):K300	x	
ClpR2 (R-ring):K226-ClpR4 (R-ring):K300	x	
ClpR2 (R-ring):K258-ClpR4 (R-ring):K300	x	

x: interlink indentified

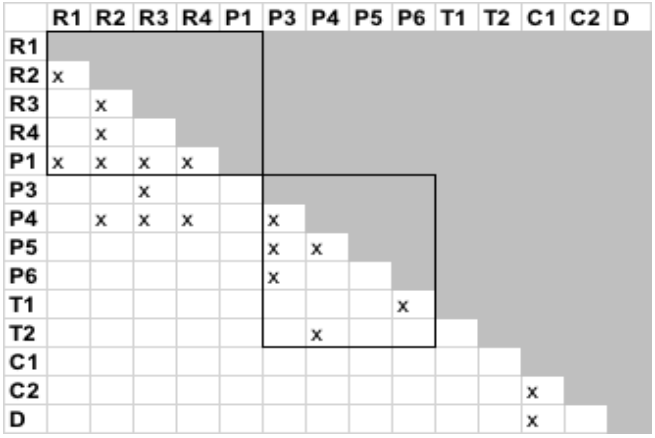


Figure 3. Proximity matrix of the CLPPRT subunits according to the inter-subunit links. x: intersubunit link(s).

responsible for protein complex assembly has been restrained during evolution. In other word, these residues are less variable and often co-evolve with the adjacent residue in 3D structure during evolution unless another biological constrain involves (mutation for certain biological functions). This approach may contribute to the investigation of the 3D structure of chloroplast CLPPRT complexes.

More and more bacterial structural studies have pointed out that AAA+ chaperones regulate the conformational state of the CLP protease through the allosteric effect caused by chaperone docking (*21-23*). To investigate the interaction between CLP chaperones and proteases, we applied the DSSO crosslinker to the affinity eluates extracted from the CLPC1E374AE718A line (wt/CLPC1E374AE718A-STREPII) (Figure 2C). The transgenic CLPC1E374AE718A-strepII line was chosen because our previous study showed that whole CLP complexes were enriched in the affinity eluate extracted from the CLPC1E374AE718A line (Montandon et al., under revision). Surprisingly, XL-MS analysis identified so far only two unique inter-subunit links within the CLPPR complex and none of the links between CLP chaperone and protease were found. Several possible reasons may explain the lack of observed inter-subunit link between CLPC1 and CLPPRT. One reason may be that the interaction was lost during sample preparation and/or affinity purification. Modified experimental conditions could help preserve the interaction between the CLP chaperones and CLP protease core complexes. Alternatively, two step crosslinking (*in vivo* mild formaldehyde crosslinking prior to cell lysis and *in vitro* DSSO crosslinking after cell lysis) may help to stabilize the transient interaction between the complexes and the

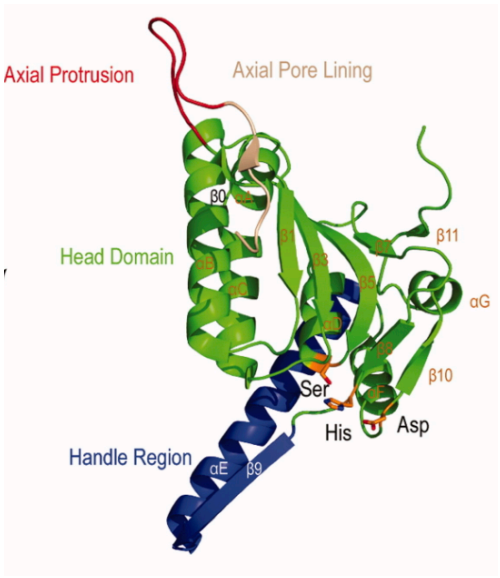


Figure 4. The crystal structure of the protomer of *E. coli* CLPP 1YG6. Adapted from Liu et al (*18*) © 2014, Taylor & Francis.

Table 2. Projection of the lysine residue on high resolution crystal structure EcCLPP (1yg6)

Subunit A	Projection to 1yg6	Sequence alignment	Subunit B	Projection to 1yg6
within R-ring				
ClpP1 (R-ring):K164	Q159	head	ClpR2 (R-ring):K266	L155
ClpP1 (R-ring):K164	Q159	head	ClpR3 (R-ring):K254	R140
ClpP1 (R-ring):K7	n/a	axial loop	ClpR1 (R-ring):K114	n/a
ClpP1 (R-ring):K7	n/a	axial loop	ClpR4 (R-ring):K282	S187
ClpR1 (R-ring):K114	n/a	N-terminal	ClpR2 (R-ring):K211	R140
ClpR2 (R-ring):K211	R140	handle (αE)	ClpR3 (R-ring):K292	E178
ClpR2 (R-ring):K211	R140	handle (αE)	ClpR4 (R-ring):K217	H122
ClpR2 (R-ring):K211	R140	handle (αE)	ClpR4 (R-ring):K264	E169
ClpR2 (R-ring):K211	R140	handle (αE)	ClpR4 (R-ring):K267	R172
ClpR2 (R-ring):K211	R140	handle (αE)	ClpR4 (R-ring):K300	n/a
ClpR2 (R-ring):K226	R140	handle (αE)	ClpR4 (R-ring):K300	n/a
ClpR2 (R-ring):K258	S187	head (β11)	ClpR4 (R-ring):K300	n/a
within P-ring				
ClpP3 (P-ring):K198	Q131	handle	ClpP4 (P-ring):K276	n/a
ClpP3 (P-ring):K198	Q131	handle	ClpP4 (P-ring):K280	n/a
ClpP3 (P-ring):K215;K218	R148; E151	handle (αE)	ClpP5 (P-ring):K273	P177
ClpP3 (P-ring):K277	n/a	C-terminal extension	ClpP6 (P-ring):K241	Q163
ClpP3 (P-ring):K289	n/a	C-terminal extension	ClpP6 (P-ring):K241	Q163
ClpP4 (P-ring):K227	R166	head (αF)	ClpT2 (wP-ring):K158	NA
ClpP4 (P-ring):K276	n/a	C-terminal extension	ClpP5 (P-ring):K276	V180
ClpP6 (P-ring):K241	Q163	head (αF)	ClpT1 (wP-ring):K139	n/a
ClpP6 (P-ring):K241	Q163	head (αF)	ClpT1 (wP-ring):K150	n/a
between P-ring and R-ring				
ClpP3 (P-ring):K288;K289	n/a	C-terminal extension	ClpR3 (R-ring):K294	V180
ClpP3 (P-ring):K288|K289	n/a	C-terminal extension	ClpR3 (R-ring):K292	E178
ClpP3 (P-ring):K288|K289	n/a	C-terminal extension	ClpR3 (R-ring):K294	V180
ClpP3 (P-ring):K289	n/a	C-terminal extension	ClpR3 (R-ring):K292	E178
ClpP3 (P-ring):K289	n/a	C-terminal extension	ClpR3 (R-ring):K294	V180
ClpP4 (P-ring):K201	R140	handle (αE)	ClpR2 (R-ring):K211	R140
ClpP4 (P-ring):K227	R166	head (αF)	ClpR2 (R-ring):K211	R140
ClpP4 (P-ring):K227	R166	head (αF)	ClpR3 (R-ring):K254	R140
ClpP4 (P-ring):K227	R166	head (αF)	ClpR4 (R-ring):K264	T168
ClpP3 (P-ring):K283;K288;K289	n/a	C-terminal extension	ClpR3 (R-ring):K294	V180

following strep purification (*9*). Another explanation for the lack of observed CLPC and CLP core interactions is simple the lack of sufficient CLPPRT complexes co-purified with the strep-tagged CLPC1E374AE718A protein in the sample. Indeed, prior analysis indicated a 5:1 protein mass ratio between ClpC1 and the CLP core complex. Given that the CLP protease contains 11 different subunits (P1, P3-6, R1-4, T1-2), the actual molar ratio between ClpC1 and each CLP core subunit is even more skewed. As mentioned previously, our first trial of XL-MS analysis using eluates extracted from ~20 g P3/P5 tissues provided little information (2 unique inter-subunit links). However, when we doubled the amount of protein, we could identify more than 20 unique inter-subunit links. This supports the idea that we simply need to increase yield to observe CLPC1-CLP protease interactions. It is noted that one of the inter-subunit links found in the XL-MS analysis of CLPC1E374AE718A sample is R4K264-P4K227. CLPR4 and CLPP4 locates in R-ring and P ring, respectively. This interlink between P-ring and R-ring implies that the CLPPRT complexes were present in the eluates. Using more starting materials to generate protein eluates may increase the inter-subunit links identified through XL-MS analysis.

REFERENCES

1. J. B. Peltier, J. Ytterberg, D. A. Liberles, P. Roepstorff, K. J. van Wijk, Identification of a 350-kDa ClpP protease complex with 10 different Clp isoforms in chloroplasts of Arabidopsis thaliana. *Journal of Biological Chemistry* **276**, 16318-16327. (2001).
2. P. D. Olinares, J. Kim, J. I. Davis, K. J. van Wijk, Subunit stoichiometry, evolution, and functional implications of an asymmetric plant plastid ClpP/R protease complex in Arabidopsis. *Plant Cell* **23**, 2348-2361 (2011).
3. L. L. Sjogren, A. K. Clarke, Assembly of the chloroplast ATP-dependent Clp protease in Arabidopsis is regulated by the ClpT accessory proteins. *Plant Cell* **23**, 322-332 (2011).
4. J. Kim *et al.*, Structures, functions, and interactions of ClpT1 and ClpT2 in the Clp protease system of Arabidopsis chloroplasts. *Plant Cell* **27**, 1477-1496 (2015).
5. J. Kim *et al.*, Subunits of the plastid ClpPR protease complex have differential contributions to embryogenesis, plastid biogenesis, and plant development in Arabidopsis. *Plant Cell* **21**, 1669-1692 (2009).
6. J. M. Holton, K. A. Frankel, The minimum crystal size needed for a complete diffraction data set. *Acta Crystallographica Section D-Biological Crystallography* **66**, 393-408 (2010).
7. D. Rovnyak, J. C. Hoch, A. S. Stern, G. Wagner, Resolution and sensitivity of high field nuclear magnetic resonance spectroscopy. *Journal of Biomolecular NMR* **30**, 1-10 (2004).
8. R. M. Glaeser, R. J. Hall, Reaching the information limit in cryo-EM of biological macromolecules: experimental aspects. *Biophysical Journal* **100**, 2331-2337 (2011).
9. C. Yu, L. Huang, Cross-linking mass spectrometry: an emerging technology for interactomics and structural biology. *Analytical Chemistry* **90**, 144-165 (2018).

10. Z. A. Chen, J. Rappsilber, Protein dynamics in solution by quantitative crosslinking/mass spectrometry. *Trends in Biochemical Sciences* **43**, 908-920 (2018).
11. A. H. Kao *et al.*, Development of a novel cross-linking strategy for fast and accurate identification of cross-linked peptides of protein complexes. *Molecular & Cellular Proteomics* **10**, (2011).
12. C. Yu, L. Huang, Cross-Linking Mass Spectrometry: An emerging technology for interactomics and structural biology. *Analytical Chemistry* **90**, 144-165 (2018).
13. A. Rudella, G. Friso, J. M. Alonso, J. R. Ecker, K. J. van Wijk, Downregulation of ClpR2 leads to reduced accumulation of the ClpPRS protease complex and defects in chloroplast biogenesis in Arabidopsis. *Plant Cell* **18**, 1704-1721 (2006).
14. F. Liu, D. T. S. Rijkers, H. Post, A. J. R. Heck, Proteome-wide profiling of protein assemblies by cross-linking mass spectrometry. *Nature Methods* **12**, 1179-+ (2015).
15. X. R. Wang *et al.*, Molecular details underlying dynamic structures and regulation of the human 26S proteasome. *Molecular & Cellular Proteomics* **16**, 840-854 (2017).
16. J. B. Peltier *et al.*, Clp protease complexes from photosynthetic and non-photosynthetic plastids and mitochondria of plants, their predicted three-dimensional structures, and functional implications. *Journal of Biological Chemistry* **279**, 4768-4781 (2004).
17. B. Zybailov *et al.*, Sorting signals, N-terminal modifications and abundance of the chloroplast proteome. *PloS One* **3**, e1994 (2008).
18. K. Liu, A. Ologbenla, W. A. Houry, Dynamics of the ClpP serine protease: A model for self-compartmentalized proteases. *Critical Reviews in Biochemistry and Molecular Biology* **49**, 400-412 (2014).
19. D. S. Marks *et al.*, Protein 3D Structure computed from evolutionary sequence variation. *PloS One* **6**, (2011).

20. T. A. Hopf *et al.*, Sequence co-evolution gives 3D contacts and structures of protein complexes. *Elife* **3**, (2014).
21. A. Martin, T. A. Baker, R. T. Sauer, Distinct static and dynamic interactions control ATPase-peptidase communication in a AAA+ protease. *Molecular Cell* **27**, 41-52 (2007).
22. M. A. Sowole, J. A. Alexopoulos, Y. Q. Cheng, J. Ortega, L. Konermann, Activation of ClpP protease by ADEP antibiotics: insights from hydrogen exchange mass spectrometry. *Journal of Molecular Biology* **425**, 4508-4519 (2013).
23. B. G. Lee, M. K. Kim, H. K. Song, Structural insights into the conformational diversity of ClpP from Bacillus subtilis. *Molecules and Cells* **32**, 589-595 (2011).

CHAPTER FIVE

CONCLUSIONS AND FUTURE PERSPECTIVE

5. 1. CONCLUSIONS

Protein homeostasis, proteostasis, is critical for cell viability and growth. Proteostasis requires the integrated activities of protein synthesis, protein folding, and protein degradation. Protein degradation by proteases needs to be well regulated since proteolysis is an irreversible reaction and costs energy. Caseinolytic protease (CLP) is central in the chloroplast protease network. Similar as bacterial CLP, chloroplast CLP is composed of multiple components (proteolytic core, chaperones, and adaptors) that fine-tune the protein degradation in response to the environmental and developmental cues. The CLP protease consists of a barrel-like tetradecamer enclosing the catalytic sites. However, chloroplast CLP protease distinguishes itself from the bacterial or mitochondria homologs by several unique structural characters: i) hetero-oligomeric composition of the tetradecamer that is associated with additional accessory proteins unique to higher plants, and ii) C-terminal extensions of the CLPP or CLPR subunits. It is unclear why chloroplast CLP has a far more diverse composition and how these different CLPP/R proteins contribute to the proteolytic system.

Here, I successfully showed that the Arabidopsis CLPP3 and CLPP5 core proteins make different contributions to the catalytic function of the chloroplast CLP machinery. Inactivation of the catalytic serine of CLPP5 (CLPP5S193A) led to

embryo lethality in the developing seeds, whereas a point mutation in CLPP3 (CLPP3S164A) did not affect plant growth and development. This phenotypic data further showed that CLPP3 makes an important structural contribution to CLP core complex since *clpp3* null mutants could not grow autotrophically, and ii) the catalytic activity of CLPP5 is essential since CLPP5S193A could not rescue the embryo lethality of *clpp5* null mutant. Whether this difference is only due to the difference in stoichiometry between CLPP3 and CLPP5 (one and three copies per complex rexpectively) within the tetradecamer, or additional structural contributions (*e.g.* additional C-terminal extension), is unclear.

In vivo trapping has been used for AAA+ proteases to systematically identify protease substrates. This approach requires substrates over-accumulating in the tetradecameric complexes prior to affinity enrichment and MS analysis. However, no obvious CLP substrate accumulation was found after extensive MS analysis of the affinity enriched CLP complexes from CLPP3/CLPP3S164A or CLPP5/CLPP5S193A lines. Only one protein (FBPA; HCEF1; AT3G54050) passed the quantitative enrichment criteria that have been used in similar experimental CLP trapping studies in bacteria. Whether FBPA/HCEF1 is a real substrate will require for further experimentation. A lack of the overaccumulation of proteins within the CLP tetradecamer with one or more inactive CLPPs suggests that proteolytic activity contributed by either CLPP3 or CLPP5 is not the bottleneck in protein substrate degradation *in vivo*. In addition, this MS data does not favor the hypothesis that each CLPP cleaves specific substrates (if so, substrate accumulation would be seen in the

MS data). Indeed, the processive nature of CLPP proteolysis argues against specific substrates for individual CLP subunits.

If indeed the individual chloroplast CLPP subunit do not have specific cleavage selectivity, whereas yet CLPPs do not show much functional redundancy, how does each CLPP and CLPR uniquely contribute to the chloroplast CLP machinery? One hypothesis is that these CLPP/R/T have co-evolved and are currently all needed for structural stability of the core complex, and/or interaction with the CLPC/D chaperones. The unique CLPPR C-terminal extensions are indeed expected to influence interaction with other proteins outside the tetradecamer. Cross-linking MS (XL-MS) combined with other structural techniques have recently developed into successful tools to determine dynamic protein-protein interaction *in vivo*. The preliminary data of the XL-MS in chapter 4 showed the possible interactions among CLPP/R/T within the tetradecamer. This preliminary work has established the pipeline of *in vitro* DSSO crosslinking of plant proteins and paves the way for a more detailed exploration of the 3D structure and the possible regulation of the chloroplast CLP machinery.

5. 2. FUTURE PERSPECTIVE

Protein degradation by the CLP system in chloroplasts must be carefully regulated. Yet its substrate recognition and degradation mechanisms are not known. Investigation of the degrons and adaptors recognized by the chloroplast CLP system will be important to understand substrate recognition, and eventually control, engineer

and tune proteostasis in chloroplasts. Tunable chloroplast protein half-life has also many potential applications in crop biotechnology. Elucidation of the structures and interactions within the chloroplast CLP system could be achieved through co-evolution analysis especially when combined with structural analysis (cryoEM, *in vivo* or *in vitro* XL, and computational modeling). Such structural elucidation should include how the individual CLP subunits and assemblies help regulate functionality of the chloroplast CLP machinery.